For cells, genome doubling brings extra benefits p. 825

Erogenous mechanoreceptors drive sexual behavior pp. 832 & 906

Tiny jets supply the solar wind pp. 833 & 867

Science

$15
25 AUGUST 2023
science.org

AAAS

CANCER CHROMOSOMES

A closer look at how aneuploidy helps tumors survive p. 848

EDITORIAL

Hawai‘i, from ashes to sustainability

Driven by fierce winds and dry, hot conditions, the fire that consumed Maui’s Lahaina—the deadliest US fire in more than a century—is a sad, stark reminder of the environmental pressures on the Hawaiian islands caused by overdevelopment and industrial tourism. As Maui emerges from the ashes, there is an opportunity to reimagine the governance of habitats through sustainable and equitable processes that preserve Hawai‘i’s natural beauty, with a less destructive, green economy that centers on ecotourism and Indigenous environmental management.

As a Native Hawaiian, I have seen the state experience several catastrophes only to have residents wonder if the next crisis will be worse. For decades, Hawaiians have warned of the islands’ vulnerability to natural disasters and human-induced stresses—from volcanic eruptions and hurricanes to a resource-consuming tourist industry that drains water tables. Yet, recovery efforts have given priority to the status quo. The devastation in Maui warrants serious changes in economic activities and infrastructures across Hawai‘i to prevent such a tragedy from happening again.

About 70% of Hawai‘i’s economy has come to rely on mass tourism. Restructuring this industry alone will require a strong vision. Visiting Hawai‘i should be an experience that deepens one’s understanding of the delicate balance of island ecosystems and the role people play in safeguarding them. An ecotourism approach directed at educating visitors about conservation and the natural environment of the islands would help build respect for nature and its vulnerabilities. This kind of tourism can be seen in highly threatened locales, including Bhutan and the Galápagos. Other progressive actions include levying higher taxes on flights and cruises to curtail the overwhelming number of visitors. Ultimately, the goal should be a gradual move toward economic models of tourism that allow more revenue to stay in the state, benefiting local initiatives, community-based tourism, and conservation.

Hawai‘i must also shift to a green economy that reduces greenhouse gas emissions and pollution. This might help mitigate some of the effects of a warming climate across the islands. This change should aim to sustainably support important businesses, including tourism, while providing new job opportunities and protecting habitats. The encouraging news is that before the Maui fire, Hawai‘i had already embarked on a transition toward renewable energy. For example, the Hawai‘i Clean Energy Initiative aims for 100% clean energy by 2045. This includes solar, wind, geothermal, and hydroelectric power and emerging technologies such as blue energy.

The knowledge of how to recover from disasters already lives in Hawai‘i, but Indigenous community voices often get lost. To understand the dire consequences of recovery without community involvement, one need only look at the aftermath of Hurricane Katrina in New Orleans. There, communities, especially those of color, were disenfranchised by practices that favored speculators. Inclusion of Indigenous expertise in environmental governance can be seen in New Zealand, Canada, and even in Hawai‘i’s recent past. In 2019, following severe flood damage, Hawai‘i capped daily visitors to Hā‘ena State Park at 900, a reduction from the previous 3000, while promoting cultural and environmental stewardship in both the park and the adjacent Nāpali Coast State Wilderness Park. This protected cultural sites and balanced local residents’ access with tourism demands. What about integrating the 2000 acres of charred remnants in Lahaina into an existing Ahupua’a, a traditional Native Hawaiian land management system? This could promote collaboration among diverse businesses, an eco-innovative mindset, and a circular economy that reduces natural resource consumption.

> “The knowledge of how to recover from disasters already lives in Hawai‘i...”

In addition to putting more Indigenous people in decision-making roles in government organizations like the Department of the Interior and Hawai‘i’s Department of Land and Natural Resources, Hawai‘i should work toward replacing traditional agencies such as the Hawai‘i Tourism Authority with Indigenous-run organizations like the Council for Native Hawaiian Advancement, ‘Āina Aloha Economic Futures, and ‘Āina Momona, which focus on blending past wisdom with forward-looking innovation.

Living in harmony with the land and sea is deeply embedded in Hawaiian traditions and practices. The future of Maui and all of Hawai‘i should be shaped by community consensus, Indigenous leadership, and ending the rampant urbanization, relentless tourism, and widespread pollution associated with environmental mismanagement. Hawai‘i can successfully recover if a new environmental governance emerges that intertwines sustainable ecosystem resilience with the profound cultural knowledge of its Indigenous communities.

–Keolu Fox

PHOTO: RILEY TAITINGFONG

Keolu Fox
is a cofounder of the Native BioData Consortium, a professor in the Department of Anthropology at the University of California, San Diego (UCSD), and codirector of the UCSD Indigenous Futures Institute.
pkfox@ucsd.edu

10.1126/science.adk4197

NEWS

> “It was abhorrent and dehumanizing work, and it was carried out under the Smithsonian’s name.”
>
> **Smithsonian Institution Secretary Lonnie Bunch,** in *The Washington Post*, about the newspaper’s reports on collections of minority group members’ remains taken without consent.

IN BRIEF

Edited by **Jeffrey Brainard**

GENOMICS

Y chromosome completely deciphered

Researchers have accomplished what was once considered an impossible task: fully sequencing not just one Y chromosome, but dozens of them from men from around the world. When a first draft of the human genome was unveiled 20 years ago, this stubby piece of DNA that makes males male proved much harder than the rest to sequence because it had so many repetitive and inverted regions. This week in *Nature*, researchers report cracking the code using newer technologies. Separately in *Nature*, a second team reports sequencing 43 more Y chromosomes collected from 21 populations. The team describes high variation in some features, such as the size of the repetitive regions. Whether this affects men in any way has yet to be determined, but one thing is clear: “The Y really is as weird, and as interesting, as we thought,” says Jenny Graves, an evolutionary geneticist at La Trobe University who was not involved in the studies.

Do advisory panels get the best?

POLICY | Research excellence is often considered a key qualification for membership on government advisory boards that help agencies evaluate scientific findings relevant to setting policy. But a study has found the probability of an individual serving on some of these panels does not correlate with measures of their scientific prowess. In fact, less-than-stellar credentials can actually improve a person’s odds of being appointed. The analysis, published this month in *Science and Public Policy*, compared the professional traits of 900 scientists at U.S. and German institutions with the likelihood of their serving on advisory panels at the U.S. Environmental Protection Agency and the German Federal Ministry of Food and Agriculture. They considered, for example, a scientist’s h-index score, a controversial metric reflecting how many papers a researcher has published and how many times these have been cited. The authors suggest the highest achieving scientists may not be motivated to serve on advisory panels because professional evaluations tend to reward scholarly papers more than this work.

RSV shot during pregnancy OK’d

PUBLIC HEALTH | The U.S. Food and Drug Administration (FDA) this week approved a vaccine given to pregnant people to protect babies from respiratory syncytial virus (RSV). The virus causes severe lung infections that are the leading cause of U.S. infant hospitalizations. Pregnant people who receive the vaccine, called Abrysvo and made by Pfizer, can pass protective antibodies to their fetuses before birth. In July, FDA approved another tool against RSV, the monoclonal antibody nirsevimab (Beyfortus), made by AstraZeneca, which is given during the first year of life.

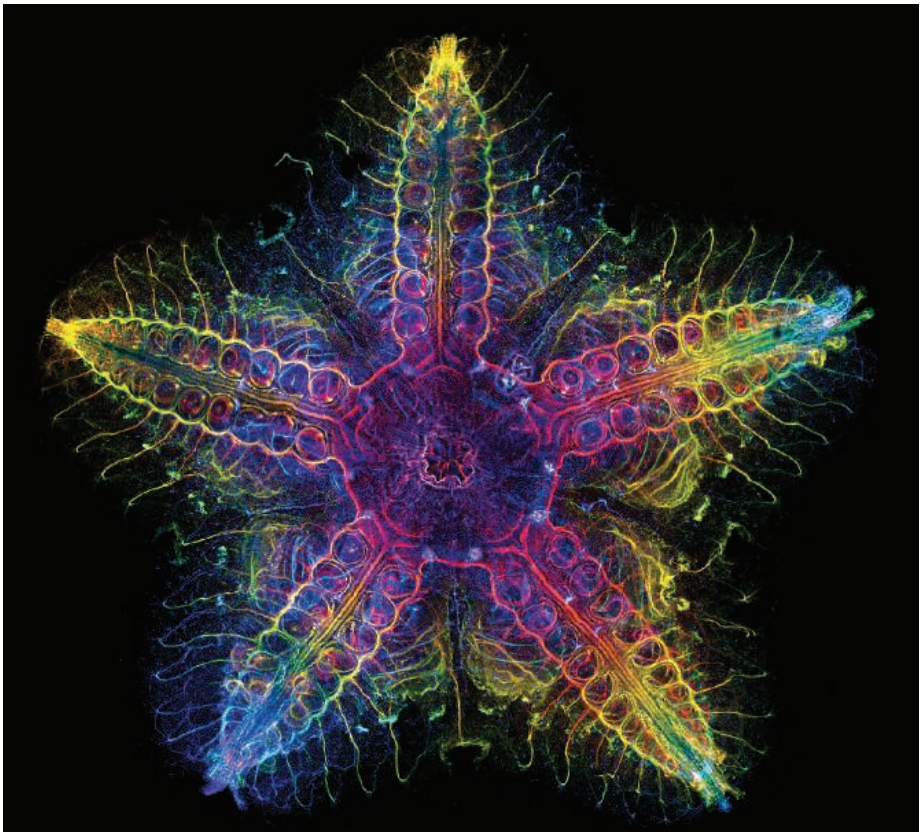
Colombia research bills blasted

ANIMAL RESEARCH | Scientists in Colombia last week persuaded lawmakers to shelve a proposal they say would have severely limited the use of animals in laboratory and field research. And they are pressing lawmakers to retract a second, similar proposal. Hundreds of scientists voiced opposition to the withdrawn bill, introduced by a member of Colombia’s ruling legislative party and backed by animal rights groups, which argued the nation needs stronger animal welfare rules. As evidence, they pointed to the recent shutdown of a primate research center in Colombia, funded by the U.S. National Institutes of Health, that allegedly mistreated monkeys. The bill’s critics said it was so broadly written it could shut down vast swaths of research, including monitoring endangered species or disease outbreaks.

Pompeii yields clues on deaths

ARCHAEOLOGY | A new examination of the famous plaster casts that record remains of people who died during the 79 C.E. eruption of Mount Vesuvius may help settle a debate about how they died and how quickly. Some experts think many residents of Pompeii, Italy, perished instantly, when very hot ash and gas evaporated their bodily fluids. Other researchers have argued that victims asphyxiated less quickly in somewhat cooler clouds of ash and gas. This week in *PLOS ONE*, archaeologists report using nondestructive x-ray fluorescence to study the chemistry of casts and bones of seven individuals found in several locations in Pompeii. That evidence, combined with other geological and anthropological clues, such as the positions of the bodies on the ground, suggest

Researchers Gianni Gallello (front) and Llorenç Alapont analyze the cast of an individual who died in Pompeii.



IN FOCUS The nervous system of a juvenile bat star (*Patiria miniata*) is highlighted in the first place image in Olympus's annual competition in life science light microscopy. Laurent Formery of the University of California, Berkeley, recorded the image after staining the animal, about 1 centimeter wide, using an antibody that binds to acetylated microtubules, structures that govern the growth of neurons. Images of echinoderms and other marine invertebrates "help communicate how much beauty we have in our oceans," Formery says, "and why it is important to know more about them and protect them."

PHOTO: LAURENT FORMERY/EVIDENT

these victims suffocated as they tried to flee the eruption and that very hot ash entombed their bodies after they died. The authors say the results indicate that people near future erupting volcanoes may have a window of time to escape, if they can protect themselves from suffocating ash clouds.

Physicist faces new scrutiny

RESEARCH INTEGRITY | Accusations of misconduct are piling up against Ranga Dias, the University of Rochester (U of R) physicist at the center of recent claims of the discovery of room temperature superconductivity. On 15 August, *Physical Review Letters* retracted a 2021 paper by Dias and nine other authors after investigations at the journal raised "serious doubts" about the authenticity of low-temperature electrical measurements reported in the paper. U of R has since said it commissioned an outside investigation. Last year, *Nature* retracted a 2020 paper by Dias and colleagues that claimed evidence for superconductivity at near–room temperature in a mixture of carbon, sulfur, and hydrogen. And this year, evidence surfaced that Dias plagiarized much of his 2013 Ph.D. thesis.

Heat could halt photosynthesis

CLIMATE SCIENCE | In a worst-case scenario of climate change, rising heat might shut down photosynthesis in tropical forests, an analysis has found. Researchers used an instrument on the International Space Station to estimate peak canopy temperatures between 2018 and 2020. The team also gathered data using ground-level sensors attached to individual leaves. Already, 0.01% of tree leaves surpass a critical temperature at which photosynthesis begins to fail, about 46.7°C. Large-scale leaf death may start to occur if warming exceeds 3.9°C, according to the study in this week's issue of *Nature*.

THREE QS

A scientist bids farewell to a troubled biodiversity spot

Cuatro Ciénegas, a basin in the Mexican state of Coahuila, is one of the most biologically diverse places on the planet. Its 300 natural pools, scattered across 800 square kilometers, host unusual microbes that Valeria Souza Saldívar, a microbial ecologist at the National Autonomous University of Mexico, has studied for almost 25 years. They include elusive Asgard archaea, organisms thought to hold clues to the evolution of more complex cells, which could inform understanding of early life on Earth and, perhaps, Mars. But farmers and others have slowly drained the pools' precious water despite pleas from researchers to stop. Souza Saldívar spoke with *Science* about her decision in June to halt her research in Cuatro Ciénegas.

Q: Why stop your work?
A: The valley is getting drier and drier. In March, we lost the Archaean Domes, [a work site] which we had been studying since 2016. The water level decreased by 25 centimeters, and all the biological richness on the surface was lost. Locals are still extracting water. There's this constant decay, and nobody does anything. Cuatro Ciénegas is dying.

Q: How does it feel to leave a site you long fought for?
A: I was very happy that I made all the waves, but leaving is a relief because it was a very heavy stone on my shoulders. Now, it's the turn of local students, nongovernmental organizations, and hotel owners [who benefit from tourists continuing to visit the pools] to make waves.

Q: What's next?
A: My Cuatro Ciénegas bacteria strains are going to Héctor Arocha Garza, director of the new science museum there, called Génesis 4C. I've started to work on the ocean. I'll start a bioremediation project, in Ohuira Bay in Sinaloa state, with a microbial consortium of seven bacteria—four of them were isolated from Cuatro Ciénegas. These bacteria work together to degrade [dissolved] nitrogen from agricultural runoff and produce nitrogen gas, instead of nitrogen oxide, which is a greenhouse gas. I believe that the lessons from the ancient sea of Cuatro Ciénegas's past must be translated into saving the sea of the future.

The world's oceans have been hit by unprecedented heat waves this year.

CLIMATE CHANGE

As oceans warm, scientists forecast heat waves

Sea temperature predictions could inform fisheries management and conservation

By **Warren Cornwall**

When heat waves began to sweep the world's oceans in June, Alistair Hobday was not surprised. The biological oceanographer had foreseen the coming temperature spikes in forecasting models he'd helped develop. The massive pool of hot water in the northeastern Atlantic Ocean, the coral-killing warmth in the Caribbean Sea, and the sweltering sea in the north Pacific Ocean had all appeared months earlier as orange and red patches on his computer screen at Australia's Commonwealth Scientific and Industrial Research Organisation (CSIRO). The forecasts, which weren't widely disseminated beyond fisheries managers and those in the fishing and aquaculture industry, proved to be a prescient warning of what was to come.

As the global climate continues to warm, scientists around the world have been working to develop models that predict when and where marine heat waves are likely to hit. CSIRO, an early leader, began to produce "experimental" forecasts for Australian waters in 2020. A separate forecast for all of the world's oceans by CSIRO and Australia's Bureau of Meteorology arrived last year. The U.S. National Oceanic and Atmospheric Administration (NOAA) unveiled its first global marine heat wave forecast in June. And Chinese scientists are working to develop predictions for their coastal waters. "I think it won't be more than a year before we've got five or six groups forecasting heat waves. This is moving really fast," Hobday says.

Scientists hope that as the models are fine-tuned, their predictions will be robust enough to alert people 3 months in advance or more, informing decisions for fisheries, aquaculture, and marine conservation. An alarming forecast, for instance, might help a regulator decide to temporarily close fishing for a heat-sensitive species.

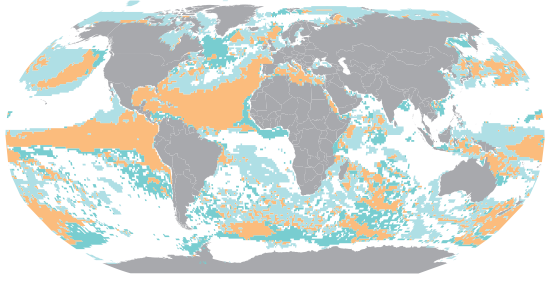
But ocean temperature predictions are not yet reliable enough to be widely embraced by managers and fishers. "The accuracy in them still needs to be worked on," says Arani Chandrapavan, a scientist in the Department of Primary Industries and Regional Development, which regulates fisheries in the state of Western Australia.

The models piggyback on existing atmosphere and ocean simulations that are already deployed to make predictions about phenomena such as droughts and the waxing and waning of the El Niño-Southern Oscillation (ENSO), a cyclical warming and cooling of water in the tropical Pacific that can influence weather around the world. "We are using [the models] to forecast extremes, specifically," says Dillon Amaya, a NOAA climate scientist and a leader in developing the agency's new forecast, which makes predictions up to 12 months in advance.

In June, his group predicted that by September half the world's ocean surface would be in a heat wave, with monthly average sea surface temperatures within the top 10% of historic

Hot water

In April, the U.S. National Oceanic and Atmospheric Administration predicted that much of the ocean surface would be engulfed in a marine heat wave (MHW) by July. Data from July matched those predictions in many areas, but not all. (An MHW was forecasted if there was a higher than 30% likelihood that sea surface temperatures would be in the top 10% of historic temperatures.)



temperatures. As of early August, heat waves stretched across 48% of the oceans, the largest swath of overheated ocean in 3 decades of satellite records. Amaya chalks that up largely to a combination of long-term heat buildup from human-caused climate change and the emergence of El Niño, the warm phase of ENSO.

Similar water temperature predictions have already been put to use in the fishing industry—in one case, to catch the same number of fish but with less effort. Fishing companies in Australia can subscribe to a 2-month forecast, created by Hobday and his colleagues, of where bluefin tuna—a temperature-sensitive species—are likely to be located off the country's southern coast. It combines a model predicting sea surface temperatures with historic tracking data for the species. The projections help the fishing boats zero in on the fish more efficiently.

In Oregon and Washington, Native tribes and state officials have consulted a seasonal forecast of regional ocean conditions over the past decade as they decide how to manage the Dungeness crab fishery. In 2018, a forecast of unusually low oxygen prompted the preemptive closure of a Washington fishery. "No one wants to pull up a crab trap full of dead crabs," says Samantha Siedlecki, an oceanographer at the University of Connecticut who led development of the forecast model, which includes temperature predictions.

Some ocean parts defy easy prediction because their temperatures fluctuate with hard-to-predict forces. The Mediterranean, which was hit by a severe heat wave last year, is strongly affected by weather on nearby land because it's relatively shallow and water moves in and out slowly from the Atlantic. And whereas ENSO influences surface temperatures in much of the world's oceans, Pacific currents near Asia aren't strongly connected to it and can change rapidly.

To test the accuracy of the NOAA model, agency scientists plugged 30 years of historical climate records into it. For many parts of the ocean, the model's forecast of whether a particular region would experience a heat wave within the coming 3.5 months matched historic reality between 85% and 93% of the time, they reported in a 2022 paper in *Nature*. But the model's predictions weren't any better than a guess in some areas, including the Mediterranean and along the Atlantic's Gulf Stream and Japan's Kuroshio Current.

Scientists have also found it challenging to predict the magnitude of temperature swings. Although the CSIRO model foresaw this summer's Atlantic heat wave near Europe, the prediction didn't show it reaching record-breaking levels. In the month of June, sea surface temperatures there reached more than 4°C above average in some places. "It's very hard to get extreme events out of models," Hobday says.

Both the Australian and U.S. forecasts are still labeled "experimental" because they haven't been deemed reliable enough to guide policymaking decisions. Still, Chandrapavan says fisheries managers in her part of Australia routinely monitor government forecasts, and that warnings of hot water have in some cases already triggered more surveys of fisheries with ship-drawn nets to check their health. Beyond a wish for greater accuracy, she and others want the forecasts to have higher resolution. They can now only predict temperatures for areas covering hundreds or thousands of square kilometers. "The big-scale forecast is like a warning. But we want to be able to operate on a spatial scale that's relevant to our stakeholders," Chandrapavan says.

She may get her wish soon. Researchers at the University of Western Australia (UWA) are embarking on an initiative funded by the government and fishing industry to develop marine heat wave forecasts at the regional level. One hope is to use artificial intelligence to produce high-resolution forecasts that demand less computer power than traditional models, says Nicole Jones, a UWA physical oceanographer leading the effort. Although such a forecast would be specific to western Australia, the tools "should be relevant worldwide," she says.

NOAA scientists also aim to bring their forecasts down to a local level in coastal waters around the United States. As part of a 4-year, $40 million initiative to help manage fisheries in a changing climate, the agency will develop models tuned to forecast ocean conditions on scales as small as 25 square kilometers, rather than the current 2500 to 10,000 square kilometers.

The hope is to be able to issue forecasts up to 1 year in advance about the increased risk of a marine heat wave in a particular area. Because conditions in the ocean often change more slowly than weather in the atmosphere, an estimate of ocean conditions one-quarter of a year into the future is often comparable in accuracy to a 2-week weather forecast, says Charlie Stock, the NOAA oceanographer who leads the initiative.

Although the result won't be a crystal ball, the information could still help guide decisions about whether to open a fishery or where to steer a fishing fleet. Stock likens it to gamblers seeking every bit of information to increase the odds of winning a five-card poker hand. Ocean temperature forecasters are now placing bets after only seeing three cards, he says. "We're trying to put out the fourth card and the fifth card so that we can kind of make those bets a little better." ■

PHYSICS

China prepares to open new neutrino observatory

Physicists hope to answer long-standing question about the relative masses of the three types of neutrinos

By **Dennis Normile**, *in Kaiping, China*

Some 700 meters beneath a hill here in the rural hinterlands of Guangdong province, a new observatory is preparing to tackle one of cosmology's lingering questions: Which of the three "flavors" of the wispy particles known as neutrinos has the greatest mass? The answer could improve our understanding of fundamental physics, astrophysics, and cosmology.

The Jiangmen Underground Neutrino Observatory (JUNO) "is very innovative and timely," says Jennifer Thomas, a particle physicist at University College London who sits on the observatory's scientific advisory committee. "It is a very challenging measurement, but if they are successful, it will have huge impact on the field." Other experiments, in the United States and Japan, are pursuing the same goal, but many physicists think JUNO's unmatched sensitivity could give it an edge.

Late last month, JUNO researchers gathered here to plan to start taking data a year from now. In blue hard hats and orange vests, some rode a rudimentary cable car down a steeply sloped 1200-meter tunnel to the experiment chamber. There, the researchers clambered up scaffolding into the center of the massive, 35-meter-diameter acrylic sphere at the heart of the detector.

The sphere will soon be filled with 20,000 tons of a liquid scintillator—an organic solution that fluoresces, or scintillates, when excited by a neutrino interacting with nuclei in its atoms. It will be the largest and most sensitive scintillation detector ever built. Scientists improved the transparency and radiopurity of the acrylic used in the sphere and of the liquid scintillator, enabling JUNO to gather more of the telltale light flashes.

temperatures. As of early August, heat waves stretched across 48% of the oceans, the largest swath of overheated ocean in 3 decades of satellite records. Amaya chalks that up largely to a combination of long-term heat buildup from human-caused climate change and the emergence of El Niño, the warm phase of ENSO.

Similar water temperature predictions have already been put to use in the fishing industry—in one case, to catch the same number of fish but with less effort. Fishing companies in Australia can subscribe to a 2-month forecast, created by Hobday and his colleagues, of where bluefin tuna—a temperature-sensitive species—are likely to be located off the country's southern coast. It combines a model predicting sea surface temperatures with historic tracking data for the species. The projections help the fishing boats zero in on the fish more efficiently.

In Oregon and Washington, Native tribes and state officials have consulted a seasonal forecast of regional ocean conditions over the past decade as they decide how to manage the Dungeness crab fishery. In 2018, a forecast of unusually low oxygen prompted the preemptive closure of a Washington fishery. "No one wants to pull up a crab trap full of dead crabs," says Samantha Siedlecki, an oceanographer at the University of Connecticut who led development of the forecast model, which includes temperature predictions.

Some ocean parts defy easy prediction because their temperatures fluctuate with hard-to-predict forces. The Mediterranean, which was hit by a severe heat wave last year, is strongly affected by weather on nearby land because it's relatively shallow and water moves in and out slowly from the Atlantic. And whereas ENSO influences surface temperatures in much of the world's oceans, Pacific currents near Asia aren't strongly connected to it and can change rapidly.

To test the accuracy of the NOAA model, agency scientists plugged 30 years of historical climate records into it. For many parts of the ocean, the model's forecast of whether a particular region would experience a heat wave within the coming 3.5 months matched historic reality between 85% and 93% of the time, they reported in a 2022 paper in *Nature*. But the model's predictions weren't any better than a guess in some areas, including the Mediterranean and along the Atlantic's Gulf Stream and Japan's Kuroshio Current.

Scientists have also found it challenging to predict the magnitude of temperature swings. Although the CSIRO model foresaw this summer's Atlantic heat wave near Europe, the prediction didn't show it reaching record-breaking levels. In the month of June, sea surface temperatures there reached more than 4°C above average in some places. "It's very hard to get extreme events out of models," Hobday says.

Both the Australian and U.S. forecasts are still labeled "experimental" because they haven't been deemed reliable enough to guide policymaking decisions. Still, Chandrapavan says fisheries managers in her part of Australia routinely monitor government forecasts, and that warnings of hot water have in some cases already triggered more surveys of fisheries with ship-drawn nets to check their health. Beyond a wish for greater accuracy, she and others want the forecasts to have higher resolution. They can now only predict temperatures for areas covering hundreds or thousands of square kilometers. "The big-scale forecast is like a warning. But we want to be able to operate on a spatial scale that's relevant to our stakeholders," Chandrapavan says.

She may get her wish soon. Researchers at the University of Western Australia (UWA) are embarking on an initiative funded by the government and fishing industry to develop marine heat wave forecasts at the regional level. One hope is to use artificial intelligence to produce high-resolution forecasts that demand less computer power than traditional models, says Nicole Jones, a UWA physical oceanographer leading the effort. Although such a forecast would be specific to western Australia, the tools "should be relevant worldwide," she says.

NOAA scientists also aim to bring their forecasts down to a local level in coastal waters around the United States. As part of a 4-year, $40 million initiative to help manage fisheries in a changing climate, the agency will develop models tuned to forecast ocean conditions on scales as small as 25 square kilometers, rather than the current 2500 to 10,000 square kilometers.

The hope is to be able to issue forecasts up to 1 year in advance about the increased risk of a marine heat wave in a particular area. Because conditions in the ocean often change more slowly than weather in the atmosphere, an estimate of ocean conditions one-quarter of a year into the future is often comparable in accuracy to a 2-week weather forecast, says Charlie Stock, the NOAA oceanographer who leads the initiative.

Although the result won't be a crystal ball, the information could still help guide decisions about whether to open a fishery or where to steer a fishing fleet. Stock likens it to gamblers seeking every bit of information to increase the odds of winning a five-card poker hand. Ocean temperature forecasters are now placing bets after only seeing three cards, he says. "We're trying to put out the fourth card and the fifth card so that we can kind of make those bets a little better." ■

PHYSICS

China prepares to open new neutrino observatory

Physicists hope to answer long-standing question about the relative masses of the three types of neutrinos

By **Dennis Normile**, *in Kaiping, China*

Some 700 meters beneath a hill here in the rural hinterlands of Guangdong province, a new observatory is preparing to tackle one of cosmology's lingering questions: Which of the three "flavors" of the wispy particles known as neutrinos has the greatest mass? The answer could improve our understanding of fundamental physics, astrophysics, and cosmology.

The Jiangmen Underground Neutrino Observatory (JUNO) "is very innovative and timely," says Jennifer Thomas, a particle physicist at University College London who sits on the observatory's scientific advisory committee. "It is a very challenging measurement, but if they are successful, it will have huge impact on the field." Other experiments, in the United States and Japan, are pursuing the same goal, but many physicists think JUNO's unmatched sensitivity could give it an edge.

Late last month, JUNO researchers gathered here to plan to start taking data a year from now. In blue hard hats and orange vests, some rode a rudimentary cable car down a steeply sloped 1200-meter tunnel to the experiment chamber. There, the researchers clambered up scaffolding into the center of the massive, 35-meter-diameter acrylic sphere at the heart of the detector.

The sphere will soon be filled with 20,000 tons of a liquid scintillator—an organic solution that fluoresces, or scintillates, when excited by a neutrino interacting with nuclei in its atoms. It will be the largest and most sensitive scintillation detector ever built. Scientists improved the transparency and radiopurity of the acrylic used in the sphere and of the liquid scintillator, enabling JUNO to gather more of the telltale light flashes.

And industrial partners boosted the sensitivity of the 43,000 photomultiplier tubes arrayed around the sphere that detect the pulses. In a first, smaller tubes sit between larger tubes, enhancing the sensors.

As neutrinos zip along at nearly the speed of light, they have the strange ability to change back and forth among three types or flavors: electron, muon, and tau. JUNO will build on results of an earlier detector, China's Daya Bay Reactor Neutrino Experiment, which used a tank holding scintillator to catch antineutrinos generated by nuclear reactors. In 2012, its data enabled physicists to determine a key parameter describing how frequently neutrinos change their identities.

Even while Daya Bay was under construction, China's Institute of High Energy Physics (IHEP) was studying a possible successor. IHEP President Yifang Wang recalls that when Daya Bay made headlines, the head of the Chinese Academy of Sciences (CAS) asked him: "What's next?" Wang "had a plan in my pocket," he says.

CAS allotted $300 million to build JUNO. To realize the plan, Wang assembled a collaboration that now numbers about 700 members from 76 institutions in 18 countries. IHEP and Wang "have leveraged their Daya Bay experience in a clever way," says Patrick Huber, a theoretical physicist at the Virginia Polytechnic Institute and State University.

The result is what Huber calls "a high-impact, modest-risk experiment." By studying electron antineutrinos coming from eight reactors at two nuclear plants, both 53 kilometers away, the project aims to solve a puzzle known as the mass ordering (also called the mass hierarchy). Neutrinos come in three mass states—m1, m2, and m3—that, confusingly, do not correspond to the three neutrino flavors. Each flavor is a quantum mechanical mix of the three mass states.

Physicists have already determined that m2 is slightly heavier than m1 and that there is a greater mass difference between m3 and the other two. But they don't know whether m3 is heavier than m2—what's called the normal ordering—or lighter than m1, the inverted ordering. The order has implications for other key questions in physics, such as how neutrinos acquire mass in the first place.

As the electron antineutrinos travel from the reactors to JUNO, many will oscillate into muon or tau antineutrinos that the detector can't see. Some of the surviving electron antineutrinos will slam into a proton in the scintillator, producing an energetic positron that results in a flash of light. The collision will also produce a lower energy neutron that, about 200 microseconds later, will be captured by an atomic nucleus and produce a second flash. The rapid one-two flicker "allows us to select our signal" from background noise, says Pedro Ochoa-Ricoux, an experimental physicist at the University of California, Irvine, who is a member of the JUNO team. By comparing the number of electron antineutrinos emanating from the reactors with the number detected, JUNO physicists can determine how many have oscillated into the undetectable muon or tau flavors.

Workers finalize the framework surrounding the Jiangmen Underground Neutrino Observatory's acrylic sphere, which will be filled with a liquid scintillator.

Whether the neutrino mass ordering is normal or inverted subtly affects the neutrino oscillations. JUNO's unprecedented sensitivity will allow scientists to distinguish whether the energy spectrum—how the neutrinos' oscillation rate varies with their energy—matches that expected for the normal or the inverted ordering. The difference between the spectra is so minute that researchers will need to record 100,000 events to achieve statistical significance, requiring about 6 years of data.

Groundwater problems delayed JUNO's start by 3 years, but Wang is confident it will solve the mass ordering question ahead of its rivals. Two ongoing neutrino experiments that study beams of neutrinos from accelerators—NOvA at Fermi National Accelerator Laboratory and Japan's Tokai to Kamioka (T2K)—"have been slowly chipping away at the mass hierarchy," Huber says. But he gives JUNO "3-to-1 odds to get there first."

If it doesn't succeed, the challenge will pass to two massive planned experiments: Japan's $800 million Hyper-Kamiokande and the United States's $3.3 billion Deep Underground Neutrino Experiment (*Science*, 30 September 2022, p. 1480). But they won't start to collect data before 2027 and sometime in the 2030s, respectively.

In the meantime, researchers expect JUNO to help answer other questions. For example, it will capture low energy solar neutrinos, providing "a real time picture of processes at the center of the Sun," says Barbara Caccianiga, a nuclear and particle physicist at Italy's National Institute for Nuclear Physics. JUNO could also score the first ever observation of the diffuse supernova neutrino background—the neutrinos emitted by all the giant stars that have collapsed in supernovae across the universe. And in 1 year JUNO will likely capture more geoneutrinos—generated by the radioactive decay of Earth's thorium and uranium—than have been measured in history. The geoneutrinos could provide clues to the planet's composition and geophysical processes such as mantle convection.

Meanwhile, "I do have another experiment after JUNO in my pocket," Wang says. It would tweak JUNO's liquid scintillator so the detector could watch for another exotic nuclear physics phenomenon called neutrinoless double beta decay. If neutrinoless double beta decay is ever observed it would indicate that the neutrino is its own antiparticle, unlike any other fundamental particle that has mass. The finding that might provide a clue to why there is more matter than antimatter in the universe. Wang aims to launch the follow-on experiment by 2040. ■

ASTRONOMY

Quasar at cosmic dawn challenges theorists

Black hole 40 million times the Sun's mass formed puzzlingly fast

Supermassive black holes power brilliant beacons by sucking in surrounding gas.

IMAGE: NASA/JPL-CALTECH

By **Daniel Clery**

Astronomers have found by far the most distant and earliest quasar ever seen, a cosmic beacon shining so soon after the big bang that standard theory can't explain how it was built.

Among the most luminous objects in the cosmos, quasars are powered by supermassive black holes in the center of galaxies, which suck in matter so voraciously that it becomes white hot from friction and glows brightly enough to be seen across the universe. Astronomers thought the black holes formed stepwise within early galaxies, as giant stars collapsed and merged, but quasars detected from when the universe was less than 1 billion years old have challenged the idea. "We were already concerned," says Anna-Christina Eilers of the Massachusetts Institute of Technology. The new one, dubbed UHZ-1, which blazed when the universe was less than 450 million years old, has made that scenario untenable.

It's not just UHZ-1's early date, confirmed in a preprint posted on 5 August. The observations show its black hole is so large compared with the galaxy around it that it can't have evolved slowly at the galaxy's heart, but must have formed rapidly, by an entirely different process.

UHZ-1 was first seen as a tiny speck of light in an image made by JWST, NASA's new infrared space telescope, of a megacluster of galaxies residing 4 billion light-years from Earth. The gravity of the giant cluster bends light like a giant lens, magnifying more distant objects behind it and making them easier to study. A few of the magnified dots appeared to be galaxies from when the universe was less than half a billion years old—among the earliest ever seen.

To see whether any of those galaxy candidates harbored quasars, observers viewed the area with another NASA space observatory, Chandra, which can detect the x-rays that are the most reliable signature of quasars. "One object stood out," says team member Andy Goulding of Princeton University. "It was booming" in x-rays.

The x-ray brightness implied that the accreting black hole had a mass roughly 40 million times the mass of our Sun. The team then went back to JWST to make sure the quasar was as early as it appeared to be in the original images. By analyzing how far certain features in its spectrum are shifted toward longer, redder wavelengths by expansion of the universe, they confirmed they were seeing the quasar at between 400 million and 450 million years after the big bang.

"It's hard to argue that it's anything other than a black hole," says Daniel Whalen of the University of Portsmouth—one far too big and too early for the conventional picture to explain. That scenario starts soon after the big bang, 13.7 billion years ago, as clumps of the mysterious dark matter that pervades the universe draw in primordial gases—mostly hydrogen and helium. The gases coalesce into giant "population III" stars, which swiftly burn up all their fuel and collapse into black holes with masses at most a few hundred times that of the Sun. Those early black holes would be the "seeds" for future supermassive ones. They would grow by accreting more gas, merge with other seeds, and pull in stars to form a galaxy around themselves.

Some theorists already doubted that this stepwise process could produce the earliest quasars, and that much larger seeds were needed. The same spectrum that confirmed UHZ-1's distance suggests they were right. It allowed astronomers to estimate the mass of the galaxy: At 140 million times that of the Sun, it was only a few times the mass of the black hole at its heart. In contrast, quasars seen later in cosmic history are dwarfed by their galaxies, being 0.1% of the mass or less. To grow so large compared with its host galaxy and so soon after the big bang, the quasar must have had a head start. UHZ-1 "doesn't put the last nail in the coffin lid" of light seeds, Whalen says. "But you really couldn't have formed this from population III stars."

In another preprint, posted on 4 August, Goulding and others argue that UHZ-1 fits a model proposed in 2017, in which radiation from early stars prevents a giant gas cloud nearby from cooling, fragmenting, and collapsing into more stars. Eventually, that cloud becomes unstable and collapses into a single black hole tens of thousands or even hundreds of thousands of times the mass of our Sun. This hefty seed could then merge with a nearby galaxy in formation and quickly grow into an outsize black hole like UHZ-1.

Another, more exotic possibility was presented in the *Proceedings of the National Academy of Sciences* last month by Katherine Freese of the University of Texas at Austin and colleagues: huge, puffy stars powered by dark matter. These "dark stars" would burn slowly, powered by the energy that some speculative forms of dark matter might give off, and over time would grow to 1 million times the mass of the Sun. When its dark matter power source gives out, a dark star would quickly collapse into a 1-million-solar-mass black hole—a perfect large seed ready to be adopted by a nearby protogalaxy. Freese says another JWST survey has already found three objects in the early universe that bear hallmarks of being dark stars.

Giant gas clouds and dark stars are only two of the scenarios theorists have devised for jump-starting the growth of quasars. Now, it's up to observers to figure out which could be real. "We need more observations and for sure they will come for such a fascinating source," says Melanie Habouzit of the Max Planck Institute for Astronomy. ■

ARTIFICIAL INTELLIGENCE

Researchers propose test for AI sentience

Human consciousness theories inform preliminary checklist for bots

By **Elizabeth Finkel**

In 2021, Google engineer Blake Lemoine made headlines—and got himself fired—when he claimed that LaMDA, the chatbot he'd been testing, was sentient. Artificial intelligence (AI) systems, especially so-called large language models such as LaMDA and ChatGPT, can certainly seem conscious. But they're trained on vast amounts of text to imitate human responses. So how can we really know?

Now, a group of 19 computer scientists, neuroscientists, and philosophers has come up with an approach: not a single definitive test, but a lengthy checklist of attributes that, together, could suggest but not prove an AI is conscious. In a 120-page discussion paper posted as a preprint this week, the researchers draw on theories of human consciousness to propose 14 criteria, and then apply them to existing AI architectures, including the type of model that powers ChatGPT.

None is likely to be conscious, they conclude. But the work offers a framework for evaluating increasingly humanlike AIs, says co-author Robert Long of the San Francisco–based nonprofit Center for AI Safety. "We're introducing a systematic methodology previously lacking."

Adeel Razi, a computational neuroscientist at Monash University and a fellow at the Canadian Institute for Advanced Research (CIFAR) who was not involved in the new paper, says that is a valuable step. "We're all starting the discussion rather than coming up with answers."

Until recently, machine consciousness was the stuff of science fiction movies such as *Ex Machina*. "When Blake Lemoine was fired from Google after being convinced by LaMDA, that marked a change," Long says. "If AIs can give the impression of consciousness, that makes it an urgent priority for scientists and philosophers to weigh in." Long and philosopher Patrick Butlin of the University of Oxford's Future of Humanity Institute organized two workshops on how to test for sentience in AI.

For one collaborator, computational neuroscientist Megan Peters at the University of California, Irvine, the issue has a moral dimension. "How do we treat an AI based on its probability of consciousness? Personally this is part of what compels me."

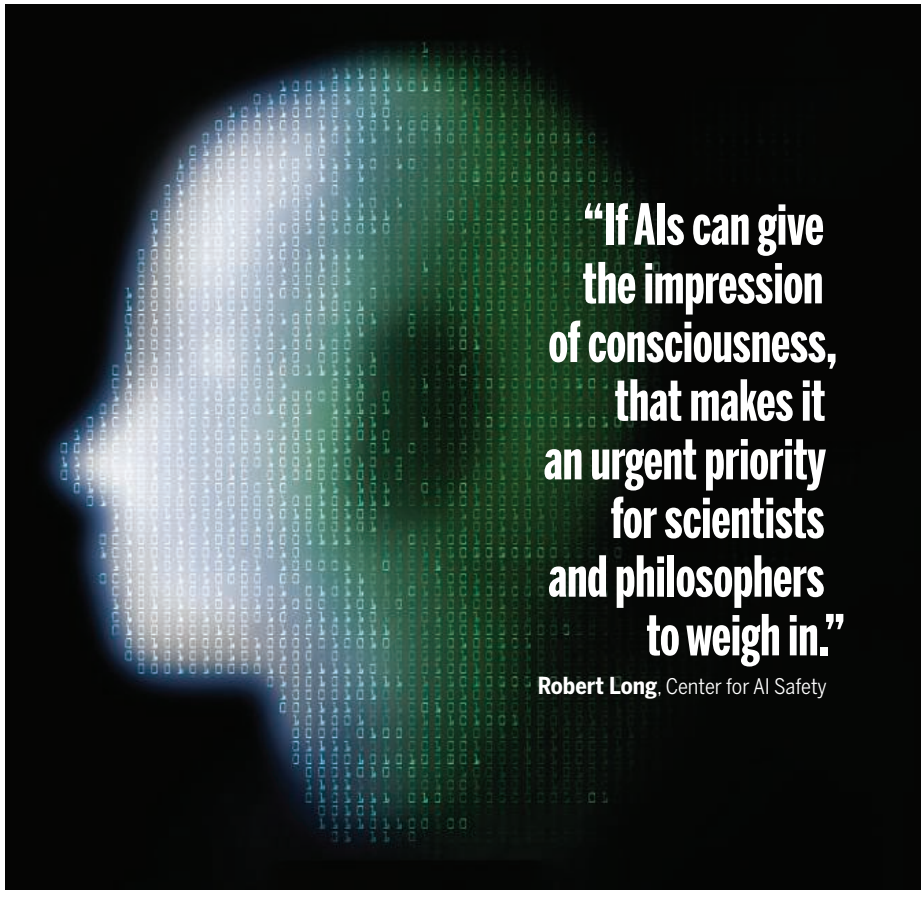
Enlisting researchers from diverse disciplines made for "a deep and nuanced exploration," she says. "Long and Butlin have done a beautiful job herding cats."

One of the first tasks for the herd was to define consciousness, "a word full of traps," says another member, machine learning pioneer Yoshua Bengio of the Mila-Quebec Artificial Intelligence Institute. The researchers decided to focus on what New York University philosopher Ned Block has termed "phenomenal consciousness," or the subjective quality of an experience—what it is like to see red or feel pain.

But how does one go about probing the phenomenal consciousness of an algorithm? Unlike a human brain, it offers no signals of its inner workings detectable with an electroencephalogram or MRI. Instead, the researchers took "a theory-heavy approach," explains collaborator Liad Mudrik, a cognitive neuroscientist at Tel Aviv University: They would first mine current theories of human consciousness for the core descriptors of a conscious state, and then look for these in an AI's underlying architecture.

To be included, a theory had to be based on neuroscience and supported by empirical evidence, such as data from brain scans during tests that manipulate consciousness using perceptual tricks. It also had to allow for the possibility that consciousness can arise regardless of whether computations are performed by biological neurons or silicon chips.

Six theories made the grade. One was the Recurrent Processing Theory, which proposes that passing information through feedback loops is key to consciousness. Another, the Global Neuronal Workspace Theory, contends that consciousness arises when independent streams of information pass through a bottleneck to combine in a workspace analogous to a computer clipboard.


"If AIs can give the impression of consciousness, that makes it an urgent priority for scientists and philosophers to weigh in."
Robert Long, Center for AI Safety

Higher Order Theories suggest consciousness involves a process of representing and annotating basic inputs received from the senses. Other theories emphasize the importance of mechanisms for controlling attention and the need for a body that gets feedback from the outside world. From the six included theories the team extracted their 14 indicators of a conscious state.

The researchers reasoned that the more indicators an AI architecture checks off, the more likely it is to possess consciousness. Mila-based machine learning expert Eric Elmoznino applied the checklist to several AIs with different architectures, including those used for image generation such as Dall-E2. Doing so required making judgment calls and navigating gray areas. Many of the architectures ticked the box for indicators from the Recurrent Processing Theory. One variant of the type of large language model underlying ChatGPT came close to also exhibiting another feature, the presence of a global workspace.

Google's PaLM-E, which receives inputs from various robotic sensors, met the criterion "agency and embodiment." And, "If you squint there's something like a workspace," Elmoznino adds.

DeepMind's transformer-based Adaptive Agent (AdA), which was trained to control an avatar in a simulated 3D space, also qualified for "agency and embodiment," even though it lacks physical sensors like PaLM-E has. Because of its spatial awareness, "AdA was the most likely … to be embodied by our standards," the authors say.

Given that none of the AIs ticked more than a handful of boxes, none is a strong candidate for consciousness, although Elmoznino says, "It would be trivial to design all these features into an AI." The reason no one has done so is "it is not clear they would be useful for tasks."

The authors say their checklist is a work in progress. And it's not the only such effort underway. Some members of the group, along with Razi, are part of a CIFAR-funded project to devise a broader consciousness test that can also be applied to organoids, animals, and newborns. They hope to produce a publication in the next few months.

The problem for all such projects, Razi says, is that current theories are based on our understanding of human consciousness. Yet consciousness may take other forms, even in our fellow mammals. "We really have no idea what it's like to be a bat," he says. "It's a limitation we cannot get rid of." ■

Elizabeth Finkel is a journalist based in Melbourne, Australia.

BIOTECHNOLOGY

Chemists convert electricity into fuel for living cells

ATP generated with renewable power could be used to manufacture proteins and medicines

By **Robert F. Service**

Power plants incessantly burn fossil fuels to convert the solar energy stored by plants eons ago into electricity. But going the other direction—converting electricity into a biologically useful form of energy—has been much more difficult. Now, however, a simple chemical scheme can convert electrical energy into adenosine triphosphate (ATP), the chemical fuel used by all cells, a research team reports. With the process, electricity from renewable sources might someday power biofactories to make everything from protein supplements to medicines.

"This is really exciting," says Michael Jewett, a bioengineer at Stanford University. "This new approach harnesses biological processes to carry out functions that nature never needed, but could benefit society."

Within the cells of plants, organelles called chloroplasts use sunlight to generate ATP as part of the process of photosynthesis. The ATP then powers myriad reactions essential for metabolism. When an ATP molecule is used, it is stripped of one of its phosphate groups, creating adenosine diphosphate, or ADP. The ADP is then recycled and fed more captured energy to regenerate ATP. Plant-eating animals burn glucose to power this same cycle, which occurs some 10 million times per second in every cell.

Industrial biotechnologists tap into this cycle by harnessing specially bred or modified microbes to manufacture everything from biofuels to pharmaceuticals. A process typically starts by growing plants to make sugar or other food that can be fed to yeast, *Escherichia coli*, or other industrial microbes. The microbes use the food to generate ATP that powers the desired biochemical reactions. But plants typically only convert 1% of the energy in sunlight into sugars or other compounds, so such processes are inefficient.

In contrast, solar cells typically convert 20% or more of the energy in sunlight to electricity. Given that disparity, Tobias Erb, a synthetic biologist at the Max Planck Institute for Terrestrial Microbiology, and his colleagues sought a way to convert electricity into ATP more directly. Others had tried before. In 2016, researchers in Spain did so by precisely orienting copies of an ATP-generating enzyme, called ATP synthase, in a membrane adjacent to an electrode. The approach worked in the lab but seemed too complex to be practical, Erb says.

Erb's team set out to devise a simpler approach, creating the "AAA cycle" in which four enzymes in solution harness electricity and use it to convert ADP, added as a reagent, into ATP. Key to the process, Erb says, is a tungsten-containing enzyme called aldehyde ferredoxin oxidoreductase (AOR) that was isolated from a bacterium just 9 years ago. AOR can't directly convert ADP to ATP. Rather it acts a bit like an engine to power the process. "AOR is an energy converter," Erb says.

AOR grabs pairs of electrons from an electrode and uses them to add an energy-rich chemical bond to a starting compound called propionate, converting it to propionaldehyde. Other enzymes then modify that chemical further, until a final enzyme remakes the starting compound, restarting the cycle while unleashing the energy in the bond. "This process releases energy that is used to generate ATP," says Shanshan Luo, a postdoctoral researcher at the Max Planck institute and lead author of the study. The team then used this ATP to drive the conversion of DNA into RNA and proteins in a cell-free setup, a result they reported last week in *Joule*.

"The simple AAA cycle is a clever and elegant approach … that is much simpler than how biology naturally makes ATP," says Drew Endy, a synthetic biologist at Stanford. It could be an enabling technology for the emerging field of electrobiosynthesis, which focuses on using electricity to power the growth of everything from food to pharmaceuticals, he adds. "It's hard for me to overstate the importance of this possibility," Endy says.

The AAA cycle must be improved, however. In solution, AOR survives only for about 1 hour. But Erb says his team is already trying to evolve more stable enzymes and protect them inside a gel they can attach to an electrode. If either succeeds, bioengineers could soon have a new way to power production processes. ■

Higher Order Theories suggest consciousness involves a process of representing and annotating basic inputs received from the senses. Other theories emphasize the importance of mechanisms for controlling attention and the need for a body that gets feedback from the outside world. From the six included theories the team extracted their 14 indicators of a conscious state.

The researchers reasoned that the more indicators an AI architecture checks off, the more likely it is to possess consciousness. Mila-based machine learning expert Eric Elmoznino applied the checklist to several AIs with different architectures, including those used for image generation such as Dall-E2. Doing so required making judgment calls and navigating gray areas. Many of the architectures ticked the box for indicators from the Recurrent Processing Theory. One variant of the type of large language model underlying ChatGPT came close to also exhibiting another feature, the presence of a global workspace.

Google's PaLM-E, which receives inputs from various robotic sensors, met the criterion "agency and embodiment." And, "If you squint there's something like a workspace," Elmoznino adds.

DeepMind's transformer-based Adaptive Agent (AdA), which was trained to control an avatar in a simulated 3D space, also qualified for "agency and embodiment," even though it lacks physical sensors like PaLM-E has. Because of its spatial awareness, "AdA was the most likely … to be embodied by our standards," the authors say.

Given that none of the AIs ticked more than a handful of boxes, none is a strong candidate for consciousness, although Elmoznino says, "It would be trivial to design all these features into an AI." The reason no one has done so is "it is not clear they would be useful for tasks."

The authors say their checklist is a work in progress. And it's not the only such effort underway. Some members of the group, along with Razi, are part of a CIFAR-funded project to devise a broader consciousness test that can also be applied to organoids, animals, and newborns. They hope to produce a publication in the next few months.

The problem for all such projects, Razi says, is that current theories are based on our understanding of human consciousness. Yet consciousness may take other forms, even in our fellow mammals. "We really have no idea what it's like to be a bat," he says. "It's a limitation we cannot get rid of." ■

Elizabeth Finkel is a journalist based in Melbourne, Australia.

BIOTECHNOLOGY

Chemists convert electricity into fuel for living cells

ATP generated with renewable power could be used to manufacture proteins and medicines

By **Robert F. Service**

Power plants incessantly burn fossil fuels to convert the solar energy stored by plants eons ago into electricity. But going the other direction—converting electricity into a biologically useful form of energy—has been much more difficult. Now, however, a simple chemical scheme can convert electrical energy into adenosine triphosphate (ATP), the chemical fuel used by all cells, a research team reports. With the process, electricity from renewable sources might someday power biofactories to make everything from protein supplements to medicines.

"This is really exciting," says Michael Jewett, a bioengineer at Stanford University. "This new approach harnesses biological processes to carry out functions that nature never needed, but could benefit society."

Within the cells of plants, organelles called chloroplasts use sunlight to generate ATP as part of the process of photosynthesis. The ATP then powers myriad reactions essential for metabolism. When an ATP molecule is used, it is stripped of one of its phosphate groups, creating adenosine diphosphate, or ADP. The ADP is then recycled and fed more captured energy to regenerate ATP. Plant-eating animals burn glucose to power this same cycle, which occurs some 10 million times per second in every cell.

Industrial biotechnologists tap into this cycle by harnessing specially bred or modified microbes to manufacture everything from biofuels to pharmaceuticals. A process typically starts by growing plants to make sugar or other food that can be fed to yeast, *Escherichia coli*, or other industrial microbes. The microbes use the food to generate ATP that powers the desired biochemical reactions. But plants typically only convert 1% of the energy in sunlight into sugars or other compounds, so such processes are inefficient.

In contrast, solar cells typically convert 20% or more of the energy in sunlight to electricity. Given that disparity, Tobias Erb, a synthetic biologist at the Max Planck Institute for Terrestrial Microbiology, and his colleagues sought a way to convert electricity into ATP more directly. Others had tried before. In 2016, researchers in Spain did so by precisely orienting copies of an ATP-generating enzyme, called ATP synthase, in a membrane adjacent to an electrode. The approach worked in the lab but seemed too complex to be practical, Erb says.

Erb's team set out to devise a simpler approach, creating the "AAA cycle" in which four enzymes in solution harness electricity and use it to convert ADP, added as a reagent, into ATP. Key to the process, Erb says, is a tungsten-containing enzyme called aldehyde ferredoxin oxidoreductase (AOR) that was isolated from a bacterium just 9 years ago. AOR can't directly convert ADP to ATP. Rather it acts a bit like an engine to power the process. "AOR is an energy converter," Erb says.

AOR grabs pairs of electrons from an electrode and uses them to add an energy-rich chemical bond to a starting compound called propionate, converting it to propionaldehyde. Other enzymes then modify that chemical further, until a final enzyme remakes the starting compound, restarting the cycle while unleashing the energy in the bond. "This process releases energy that is used to generate ATP," says Shanshan Luo, a postdoctoral researcher at the Max Planck institute and lead author of the study. The team then used this ATP to drive the conversion of DNA into RNA and proteins in a cell-free setup, a result they reported last week in *Joule*.

"The simple AAA cycle is a clever and elegant approach … that is much simpler than how biology naturally makes ATP," says Drew Endy, a synthetic biologist at Stanford. It could be an enabling technology for the emerging field of electrobiosynthesis, which focuses on using electricity to power the growth of everything from food to pharmaceuticals, he adds. "It's hard for me to overstate the importance of this possibility," Endy says.

The AAA cycle must be improved, however. In solution, AOR survives only for about 1 hour. But Erb says his team is already trying to evolve more stable enzymes and protect them inside a gel they can attach to an electrode. If either succeeds, bioengineers could soon have a new way to power production processes. ■

ASTRONOMY

Cyberattack shutters major telescopes

Denied remote access, astronomers struggling to prevent loss of valuable observing time

By **Celina Zhao** *and* **Tanvi Dutta Gupta**

A mysterious "cyber incident" has paralyzed 10 telescopes—including major instruments in Hawaii and Chile—since the beginning of August and allowed only in-person observations at several more. As *Science* went to press, research teams were scrambling to find alternatives, including sending graduate students to Chile to relieve exhausted on-site staff, as critical observation windows threaten to close.

"We're all in this together," says Gautham Narayan, an astronomer at the University of Illinois Urbana-Champaign whose team is trying to save its chance to observe new supernovae using one of the affected Chilean telescopes. The astronomy community has a "grim determination to press on, despite the trying circumstances," he adds.

Normally, computer systems operated by NOIRLab, a ground-based astronomy center funded by the U.S. National Science Foundation (NSF), allows researchers to remotely operate a far-flung network of telescopes. But on 1 August, NOIRLab announced the detection of an apparent cyberattack on its Gemini North telescope on the Big Island of Hawaii. Whatever happened may have placed the instrument in physical jeopardy. "Quick reactions by the NOIRLab cyber security team and observing teams prevented damage to the observatory," NOIRLab said.

NOIRLab powered down all operations at the International Gemini Observatory, which runs the Hawaii telescope and its twin, Gemini South, on Cerro Tololo in Chile. (The latter was already offline for a planned outage.) On 9 August the center announced it also disabled remote observations at a number of other telescopes in Chile, including the Víctor M. Blanco 4-meter and Southern Astrophysical Research (SOAR) telescopes.

NOIRLab has provided few further details, even to employees, and declined to say whether the incident was a ransomware attack, in which hackers demand money for restoring data or a network. NOIRLab was "working around the clock to get the telescopes back into the sky," a spokesperson told *Science* last week.

The Gemini North telescope in Hawaii is one of a number of instruments affected by a "cyber incident."

Narayan praises NOIRLab's "exemplary" response, and he and other astronomers express sympathy for the center. "I assume the challenges they're facing are bigger than me not getting observations," says Luis Welbanks, an astronomy postdoc at Arizona State University.

But astronomers are getting anxious. Multiple international projects, as well as doctoral theses and papers under development, depend on data from the telescopes.

One pressing problem is that ground-based astronomy often depends on observations precisely timed to when extraterrestrial objects align with the field of view of a specific telescope. Astronomers try to plan for delays—anything from bad weather to a power outage—but hackers have not typically figured into their calculations. "We're lucky enough to make it through a regular night," Welbanks says. "But now we have to consider the cybersecurity implications."

For early-career researchers like Welbanks, a lengthy delay could be particularly harmful. He relies on Gemini South to study exoplanet atmospheres; the shutdown has already caused him to miss three of his seven observation windows this year.

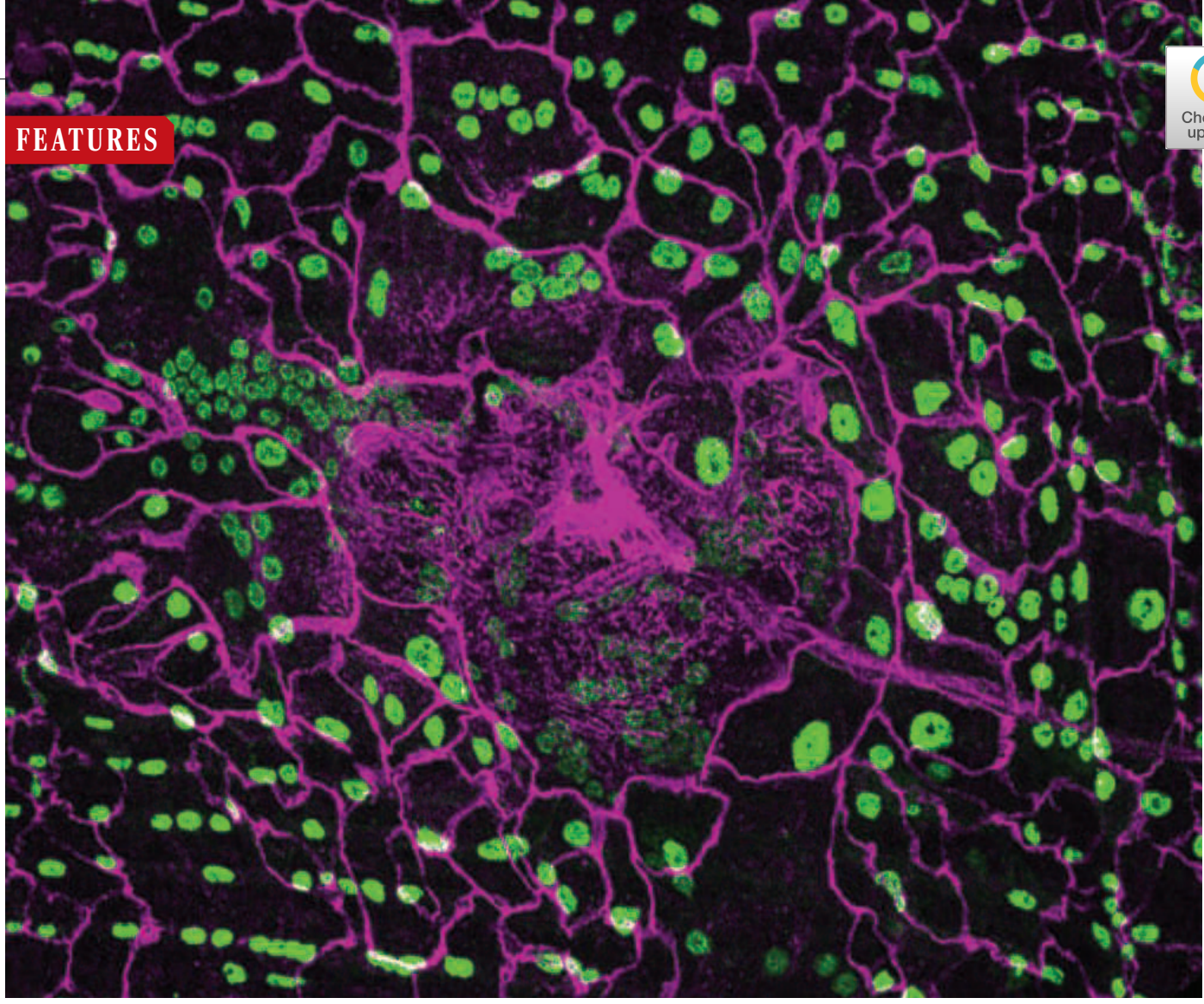
NOIRLab staff have been going "well above and beyond the call of duty" to keep projects going, Narayan says. To make up for the lack of remote observing, some on-site staff at the Blanco and SOAR telescopes have been helping researchers by recording in-person observations. But that practice is not sustainable, NOIRLab noted in an internal email—hence the discussions about dispatching graduate students to Chile.

Cybersecurity experts are perplexed as to why the hacker targeted Gemini North. "Quite possibly, the attacker doesn't even know they are attacking an observatory," says Von Welch, retired lead of the NSF Cybersecurity Center of Excellence.

The episode is another wake-up call for the astronomy community, Welch and others say. In November 2022, a cyberattack forced the Atacama Large Millimeter Array radio telescope in Chile to go dark for nearly 2 months. Highly collaborative research institutions such as NOIRLab face formidable security challenges, Welch notes. Unlike private companies, for example, they can't easily wall off their systems. Erect high barriers, Welch says, and "you just broke all the scientific workflows."

Despite the lack of clarity over how the Gemini North and NOIRLab systems were compromised, astronomers say the attack provides motivation to improve cybersecurity practices. The astronomical community needs to rethink how it manages network access, Narayan says, and understand how damaging something as simple as a lost password can be.

"The strongest, most impenetrable fortress in the world" won't help astronomers "if you forget to lock even a single door or window," says Patrick Lin, who leads an NSF-funded space cybersecurity grant at California Polytechnic State University. "The weakest link is often with us, the humans." ■



To heal a fruit fly's abdominal wound (center), some cells (pink outlines) fuse and become giants, containing many nuclei (green).

IMAGE: LOSICK LAB/BOSTON COLLEGE

STRESS RESPONDERS

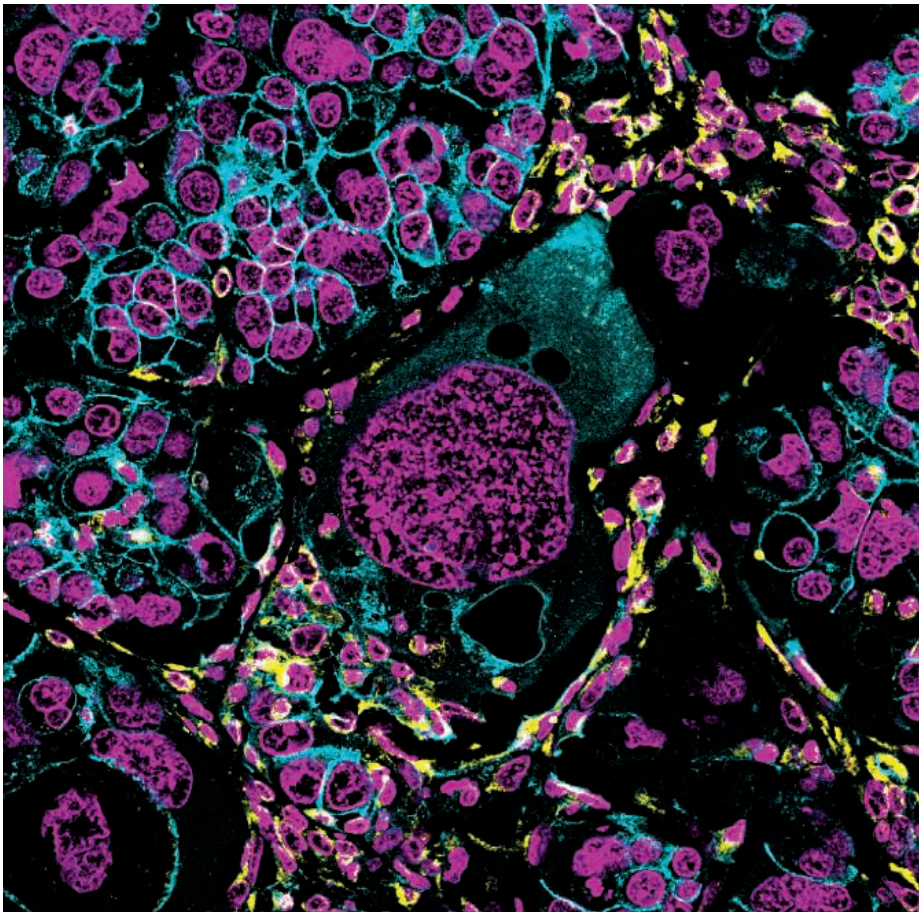
Polyploid cells, which have extra copies of their genomes, may help tissues respond to injuries and species survive cataclysms

By **Elizabeth Pennisi**

When Vicki Losick got her Ph.D. and joined a fruit fly lab at the Carnegie Institution for Science in 2008, its head announced that he expected his postdocs to launch new fields of inquiry. She chose a then-fashionable focus: stem cells, versatile cells that specialize into other cell types and play critical roles in embryonic development and the renewal of adult tissues. Losick wondered whether they also help in wound repair. So, she and another postdoc, Don Fox, began stabbing fruit flies with a tiny needle, hoping to document stem cells coming to the rescue.

Instead, the two postdocs, working independently, saw other cells near the wounds behaving oddly. The cells grew and prepared to divide by duplicating their DNA. Then they stalled, each remaining a single, enlarged cell with multiple copies of its genome. "I was shocked," recalls Losick, now at Boston College.

When she and Fox looked at the fly wound sites a few days later, they saw signs



The large nucleus (pink, center) of a giant cancer cell has multiple sets of chromosomes that help the tumor survive treatment.

that these so-called polyploid cells, and not stem cells, were the major wound healers. At the puncture site, supersize cells with multiple nuclei quickly closed up the wound. "Simultaneously we found the same thing [and it] had nothing to do with stem cells," Fox recalls.

Since then, the two scientists' fascination with polyploid cells has only grown. Most plant and animal cells are diploid—their genomes consist of two sets of chromosomes, one from each parent. Polyploid cells, which have one or more additional sets of chromosomes, were once thought to be harmful anomalies, notorious for helping drive cancer growth (*Science*, 14 February 2014, p. 725). But it's now clear, Losick says, that polyploidy "is widespread in the fruit fly and I think in our body as well." Indeed, some data suggest 80% of the cells in the human heart are polyploid by adulthood, up from none at birth.

As polyploid cells emerge as common and seemingly crucial, this once obscure topic is now bringing together cancer researchers, developmental biologists, evolutionary biologists, cell biologists, and agricultural scientists. About 150 researchers gathered in Florida in May for a meeting Losick helped organize, Polyploidy Across the Tree of Life, trading information from their normally siloed fields. "The amount of cross-interest was just mind-boggling," says Shyama Nandakumar, a developmental biologist at the University of Pittsburgh.

The research effort has revealed several genes that naturally help drive polyploidy, evidence that it is not just a cellular accident. Researchers have found clues that polyploid cells shape the development of tissues from plant cells to hearts (see sidebar, p. 828). And they are amassing evidence that polyploid cells are an essential response for all sorts of tissues and organs coping with stress, from injuries like those in the fruit flies to disease and damage in human lungs, livers, and kidneys. Polyploid cells "are the stem cell alternative," suggests Fox, now at Duke University.

Polyploidy may be a coping method not just for individual organisms, but for entire species. About 30% of plants are fully polyploid, meaning all their cells have double or more the number of chromosomes found in their ancestors. The same is true for a small number of animals—several salamanders, for example. Evolutionary biologists have found that polyploid organisms are often at a competitive disadvantage compared with diploid counterparts, raising questions about why the trait persists. By pinpointing when genome duplications occurred millions of years ago, researchers are glimpsing a potential answer: Polyploidy may help species weather catastrophic environmental changes. At all sorts of spatial and temporal scales, polyploidy is a "damage response," suggests Douglas Soltis, a plant evolutionary biologist at the University of Florida. "It's the single most important evolutionary thing that nobody talks about."

FOR ALMOST A CENTURY, scientists have recognized that polyploid cells can arise in multiple ways (see graphic, p. 827). Sometimes cell division is interrupted midstream, and the replicated DNA gets stuck in the parent cell instead of being divided between daughter cells. In other scenarios, cells fuse to make a single cell with multiple nuclei. In each case, the result is bigger cells. "By becoming polyploid, the [cells] change their physical properties," Losick says.

The cells are sometimes harmful, helping cancers develop or persist despite chemotherapy and radiation. Polyploid cells can also damage the kidneys and liver under some conditions. But Losick and Fox glimpsed another role in their study of the fruit fly wounds. When they surveyed the scientific literature for similar finds, they learned that other groups had seen polyploid cells appear in diseased or stressed tissues, including in livers after the organs were injured. So, the pair investigated its fruit fly system further. They found the number of polyploid cells and the number of nuclei and genomes in individual cells varied depending on factors including the size of their wound.

Most important, if they inhibited the creation of polyploid cells by tinkering with genes needed for chromosome doubling or cell fusion, wounds failed to heal. Even before the Carnegie group reported the finding in 2013, Wu-Min Deng, a fruit fly biologist at Tulane University, reported the same phenomenon in wounded fruit fly ovaries. The publications signaled "a new frontier in regeneration biology," Fox says.

Fox, Losick, and others also documented how the cells contribute to healing. In fruit flies, some of the mammoth cells quickly plug the space left by puncture wounds, and because they have multiple copies of various genes, they can make extra proteins that speed healing. For example, where the needle poke destroys muscle, these cells make a lot of myosin protein, which helps muscles contract. The extra

DNA of polyploid cells, Losick suggests, may also make them resistant to injury-induced DNA damage from inflammation that would cause a normal diploid cell to die, bogging down recovery.

A similar process unfolds in mouse kidneys injured by toxins and in human kidneys subjected to dehydration, trauma, infection or other stresses. At the May meeting, University of Florence cell biologist Letizia De Chiara reported that she and Florence nephrologist Paola Romagnani observed new, big polyploid cells quickly engulfing injured areas, restoring kidney function in the mice. She and colleagues also looked at kidney biopsies of people in whom the organ was gradually failing. Many polyploid cells filled those kidneys. In people who had recovered from acute kidney injury, the number of such cells declined over time, she reported at the meeting.

Once its damage control work is done, polyploidy has to shut down because the accumulation of these supersize cells can be harmful. The Florence team reported last year, for example, that they can lead to scarring and chronic kidney failure. But Losick's graduate student Loiselle Gonzalez has found that certain kinds of scarring itself can help control their numbers. When she interrupted scarring in wounded flies, polyploid cells continued to form and their wounds never fully closed up, she reported at the May meeting. "Scarlike tissues, including fibrosis, may be necessary to constrict polyploidy," Losick says.

Losick has identified a protein that helps control polyploidy in fruit flies: the fly equivalent of a mammalian molecule called YAP1, for yes1 associated transcriptional regulator. YAP1 was known to help regulate genes that control organ size. It also turns out to stimulate polyploidy in insect wound healing, and to curb the process when its activity falls. The Florence team found YAP1 has the same role in mouse kidneys. "It was amazing that we were seeing very similar things, in the *Drosophila* [abdomen] and in the mammalian kidney," Romagnani says.

The mouse studies suggest that inhibiting YAP1 at the right time might make damaged kidneys less likely to develop harmful scarring, De Chiara says. She and her colleagues are investigating this pos-

Chromosomes gone wild

Cells in most species have two matching sets of a well-defined number (N) of chromosomes. But several processes can endow normally 2N cells with extra sets. The condition, known as polyploidy, can change cells' appearance, give them new properties, and sometimes lead to new species.

Nuclear expansion

When a cell begins to divide, each chromosome duplicates, forming linked chromatids. They normally split and separate into two cells. But sometimes the separation fails, resulting in a single cell with 4N and subsequently 8N chromosomes in a bigger nucleus.

Nuclear multiplication

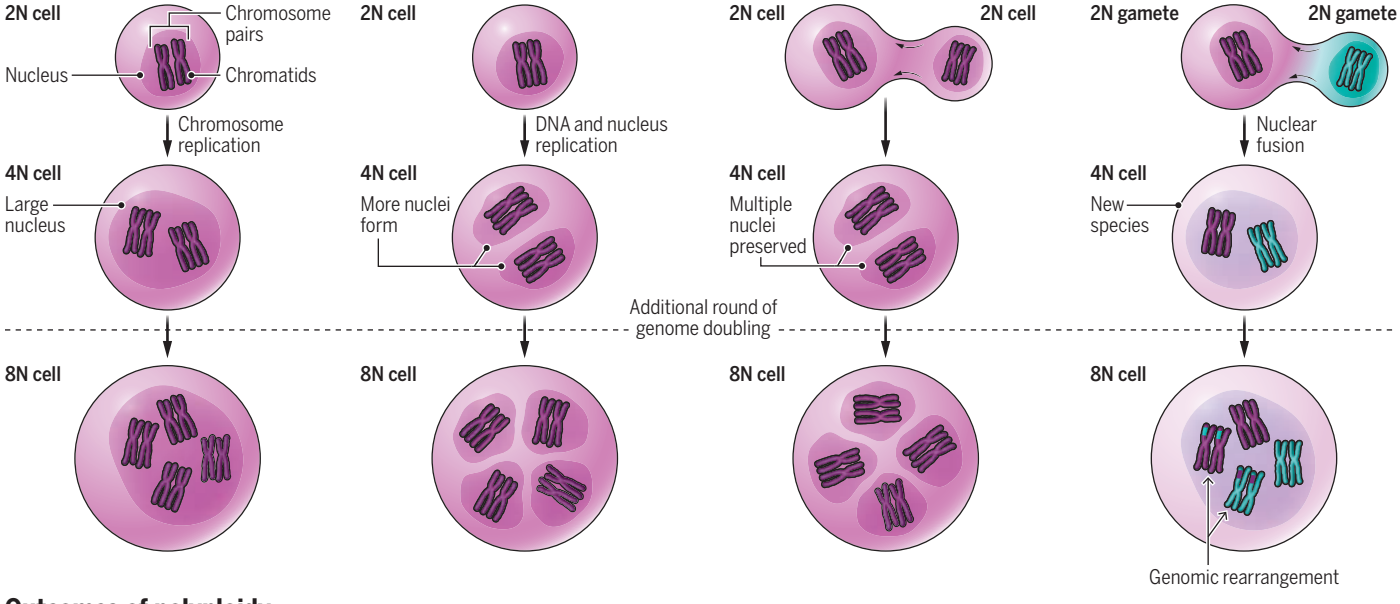
Sometimes a cell starts to replicate and new nuclei form, each with a normal chromosome complement, but the cell fails to divide. Each nucleus is normal size but the cell itself expands and the total chromosome number can double or quadruple.

Cellular fusion

Cells in trouble can save themselves by fusing, resulting in a larger cell with multiple chromosome sets. The resulting polyploidy can protect against cell death but also cause genomic rearrangements.

Making a new species

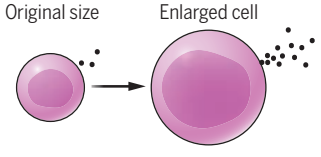
Reproductive cells, or gametes, sometimes have 2N instead of the usual N chromosomes. Especially in plants, 2N gametes from the same or different species can fuse and form new species.



Outcomes of polyploidy

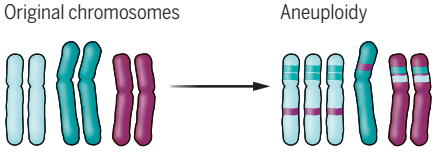
More capable cells

The extra DNA that swells polyploid cells also spurs protein production and metabolic activity.



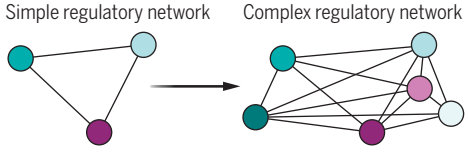
Genome shock

Additional chromosome sets can perturb, or shock, DNA replication, leading to aneuploidy—extra, missing, or rearranged chromosomes.

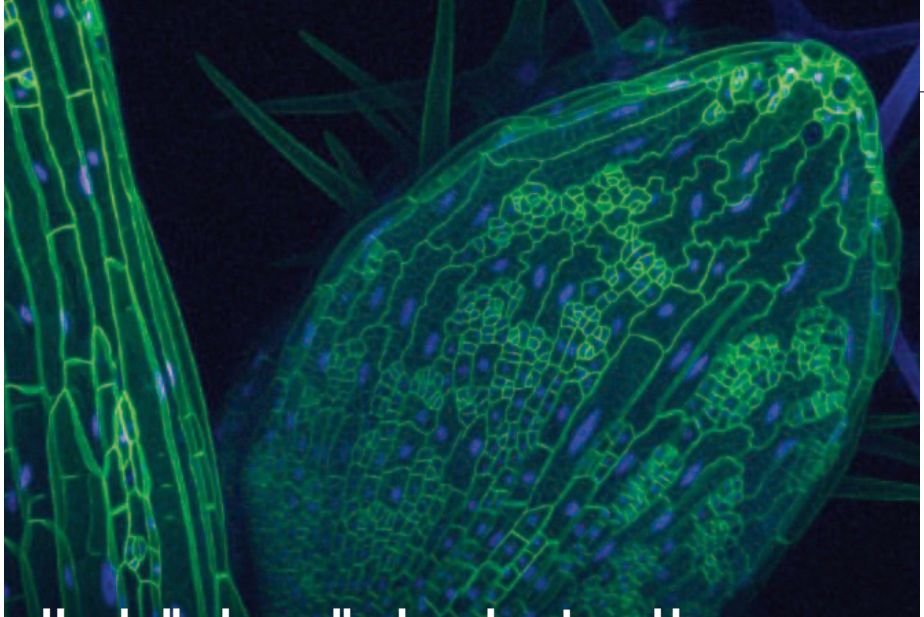


More complex gene networks

Duplicated DNA means each gene has exponentially more regulatory connections to other genes, fine-tuning the cell's gene activity.



GRAPHIC: A. MASTIN/*SCIENCE*



How bulked-up cells shape hearts and leaves

Giant cells filled with extra chromosomes may deserve credit for a pleasurable springtime sight: green buds unfurling to reveal brilliant blossoms. Similar so-called polyploid cells appear to help tissues and even whole organisms recover from injury and handle other types of stress (see main story, p. 825). But there's increasing evidence that these cells, which have more than the two sets of chromosomes found in ordinary diploid cells, are also an essential part of development in both plants and animals, their large size helping sculpt flowers, leaves, and even the developing heart.

Plant biologists have long known that sepals, the modified leaves that surround a developing flower, contain some of these giant cells. A decade ago, Cornell University's Adrienne Roeder showed that a precise number of polyploid cells is key to proper flowering. By mutating genes that control polyploidy in the mustardlike plant *Arabidopsis thaliana*, she and her team discovered that if the sepal had too few polyploid cells, it was stiff and upright and blooming was impeded. Too many such cells caused the sepals to splay and expose the petals prematurely. "You need just the right amount [of polyploidy] to get the right amount of curvature," says Don Fox, a developmental biologist at Duke University.

More recent work has unearthed a key molecular signal driving polyploidy in sepals. Roeder and her postdoc Batthula Vijaya Lakshmi Vadde discovered that normal, diploid cells turn into polyploid versions when a protein, ATML1, kickstarts the process, in part by stimulating the production of certain fatty acids, And Roeder's graduate student Frances Clark has found that polyploidy, spurred by the same molecular signals, also shapes ordinary leaves. In *Arabidopsis*, the first giant cells develop at a leaf's tip when it is about 1 week old. Then more such cells appear ever closer to the leaf base, forming on both the top and bottom of the leaf, Clark reported. When plants are engineered to produce more polyploid giant cells, the leaf becomes more oblong.

Meanwhile, Fox's postdoc Archan Chakraborty has found that the tube-shaped heart of the fruit fly *Drosophila* is rich in polyploid cells. During development cells in the larger, rear part of the heart become more polyploid—with more sets of chromosomes—and bigger than those in the front part. Chakraborty suspected the outsize cells help bulk up that part of the fly's heart, the equivalent of the human heart's muscular ventricles. To test the idea, he reduced flies' levels of insulin, which lowered the number of polyploid cells in the heart. The result: The flies entered the equivalent of heart failure, he reported at a May meeting and on 1 August in *Development*.

By examining donated human hearts, Fox and Chakraborty have shown that ventricles are also more sensitive to insulin and the cells there have more sets of chromosomes than cells in the atria, the heart's smaller chambers. To Wu-Min Deng, a fruit fly biologist at Tulane University, the finding suggests that people, too, may owe their healthy hearts to overstuffed cells. "A conserved mechanism may exist in human heart development," he says. —*E.P.*

In young *Arabidopsis* leaves treated to develop more polyploidy, many cells (green outlines) become supersize and contain extra-large nuclei (blue).

sibility in animal models. Several biotech companies are developing drugs that target the YAP1 pathway for other purposes, and they might also help defend against the downsides of polyploidy.

Drugs targeting polyploid cells might also help counter one of their most destructive effects: their ability to help cancer cells withstand the stresses of treatment. In most tumors, at least a few cells transform into nondividing polyploid giant cells, whose extra genetic material makes them resistant to chemotherapy or radiation. These cells also seem to have physical properties—increased motility and elasticity—that improve their ability to leave the tumor site and colonize other parts of the body, researchers from the Johns Hopkins University School of Medicine reported last year on bioRxiv.

Another Hopkins team found a multiprotein complex that seems to play a key role in the cells' resilience. When they inhibited one component of the complex, a protein called CDK9, in the polyploid cells, the cells no longer survived exposure to anticancer drug—suggesting a tempting vulnerability in drug resistant cancers, they reported at the meeting.

IF POLYPLOID CELLS are welcome short-term visitors in damaged or stressed tissues but a menace when they linger in cancers, why does polyploidy persist for many generations—and in nearly every cell—in some species including crops such as wheat, strawberries, and sugar cane?
In most of these species' past, two or more relatives interbred, and their genomes joined forces instead of reverting to a diploid. Sometimes these joined genomes duplicated yet again—wheat has six sets of chromosomes; strawberries have eight. The multiple copies of key genes for growth and nutrients can result in denser heads of grain, bigger fruit, and taller, sweeter stalks.

Those traits may delight human consumers, but for plants, "many, if not most, whole genome duplications are actually maladaptive," says Yves Van de Peer, an evolutionary biologist at Ghent University. Polyploid plants need more nutrients, for example, and seem to grow slower than their diploid counterparts. Yet the trait persists, and Van de Peer believes the answer, again, comes down to coping with stress.

He and his colleagues began to pin down when wild plant species' genomes duplicated to see whether there was any pattern. They found many of these duplications occurred about 66 million years ago, right around when an asteroid hit Earth and caused a mass extinction. The trend held up

in a further analysis of dozens more plant genomes, which revealed additional waves of whole genome duplication during periods of climate change or glaciation.

"I was kind of skeptical at first [about Van de Peer's idea]," Soltis says. "But it continues to stand the test of time." Van de Peer's team found, for example, that some species of *Arabidopsis*, a genus often used in plant biology studies, became polyploid during cold spells over the past 2 million years. In 2019, Harvard University evolutionary biologist Charles Davis concluded that a group of tropical flowering plants called Malpighiales underwent whole-genome duplications during an intense warming period 50 million years ago. Today they comprise some 16,000 species, suggesting the polyploidy has helped more than hurt.

Van de Peer now believes the increased genetic flexibility provided by multiple genomes lets polyploids make quick adjustments to new stresses and ride out catastrophic events that wipe out most normal plants and animals. The pattern of whole genome duplications "suggests a short-term, perhaps even immediate, evolutionary advantage for polyploids," Van de Peer says.

At the meeting, he offered computer modeling and experimental results that back up that idea. His team created swarms of digital organisms, endowed with "genes" that activate, say, virtual wheels and make the organism move right, go a certain speed, or avoid or approach other organisms.

In some swarms, individuals had one copy of a "genome" and in others, two. To test the fitness of the two types, Van de Peer assessed how well the individuals move along a grid to find digital "food." Those with one genome did much better until the researchers made food scarce. "Then the polyploids started to show more extreme behavior," Van de Peer says. They took bigger steps, found more food, and even worked together to prey on their competitors. Eventually the one-genome swarm went extinct. The work "confirms that under a stressful environment, having a duplicated genome was a good thing," he reported in 2019 in *PLOS ONE*.

"The fact that they can outcompete the diploid versions is kind of a remarkable experiment," Soltis says. "There's definitely something going on."

Now, Van de Peer believes he knows what. The genes in the simulation formed interacting networks, and Van de Peer found that in the organisms with double the genes, the networks wound up with many more connections. They made it possible for the organisms to move faster, jump farther, and make unexpected moves or interact with other organisms in more complex ways.

Something similar is likely to happen in living organisms, he and others believe. "Because you have more chromosomes, you have more knobs to turn to fine-tune the dosage of genes relative to each other," Fox explains. That leads to more of certain proteins and less of others and can change the organism's behavior, physiology, or chemistry. Polyploid species "are primed to adapt to whatever the environment throws at them," Losick adds.

Much of that fine-tuning might not be needed for, say, a stable environment in which the diploid organism is already well adapted. But when the dinosaur-killing asteroid hit, darkening Earth with smoke from wildfires, polyploids may have had the versatility to survive, Van de Peer proposes. "Usually, whole-genome duplication is an evolutionary dead end," he concludes. "But if they occur at the right time, they can create evolutionary opportunities."

To test the idea, Van de Peer and colleagues have turned to a small aquatic plant called greater duckweed (*Spirodela polyrhiza*). They induce polyploidy in some individuals by exposing them to a chemical that interrupts cell division. The researchers then grow the diploid and polyploid versions side by side to track how they cope with stresses such as high salinity or large concentrations of heavy metals. The polyploid plants proved more resilient, he reported at the May meeting. Soon, his team will start to sequence duckweeds from different generations and assess their gene activity to see how gene regulatory networks might differ between the polyploids and diploids.

Tia-Lynn Ashman and Martin Turcotte, evolutionary ecologists at the University of Pittsburgh, have also studied polyploidy in duckweed populations. The polyploid plants tended to grow more slowly and reached smaller population sizes than their diploid counterparts. But they supported a more diverse ecosystem of microbes, Ashman reported at the meeting.

That could be another beneficial consequence of polyploidy. Increasingly, researchers are recognizing that an organism's microbiome helps it survive, so a more diverse roster of microbial guests may enable the host to digest more kinds of food or enhance resilience in other ways. The increased diversity "could provide a mechanism for the broader ecological range of polyploids seen at the global scale," Ashman suggests.

When duckweed is forced to become polyploid, its cells, and consequently its leaves, become much bigger (right) than normal (left).

PHOTO: AUDREY BURR

What is certain is the polyploid cells, far from being abnormal, are one of life's major mechanisms for coping with the stresses of injury, disease, and a hostile environment. At the meeting, "There was the increasing realization that whole-genome doubling does not simply double everything in the cell—rather we are seeing unique biology," says Pamela Soltis, a plant evolutionary biologist at the Florida Museum of Natural History.

That realization pleases Losick, who recalls the high expectations of her postdoc adviser. "I feel proud," she says—"excited to be part of this new field." ■

CONSERVATION

Forest carbon offsets are failing

Analysis reveals emission reductions from forest conservation have been overestimated

By **Julia P. G. Jones**[1,2] *and* **Simon L. Lewis**[3,4]

Conserving tropical forests is of utmost importance for the future of humanity and biodiversity. Changes in land use, mostly deforestation in the tropics, emit 5 billion metric tons of carbon dioxide annually—second only to fossil fuel use, which emits 35 billion tons (*1*). Reducing emissions to net zero is necessary to stabilize global temperatures (*2*). One controversial approach to tackle fossil-fuel emissions from private companies, individuals, and governments has been to "offset" them by investing in projects to either stop emissions that would have otherwise occurred, such as by reducing deforestation, or by investing in carbon uptake projects, such as forest restoration. On page 873 of this issue, West *et al.* (*3*) show that offsetting through paying projects to reduce emissions by conserving tropical forests is not reducing deforestation as claimed and is worsening climate change.

West *et al.* studied 26 projects spanning six countries across three continents. These projects issued REDD+ (Reducing Emissions from Deforestation and Degradation) credits to the voluntary carbon market. Each credit equates to 1 metric ton of carbon dioxide that has not been emitted because of the existence of the project, which conducts activities to lower deforestation in the project area. The credits studied by West *et al.* were issued under the Verified Carbon Standard, the largest crediting program in the voluntary carbon

[1]College of Environmental Science and Engineering, Bangor University, Bangor, UK. [2]Department of Biology, Utrecht University, Utrecht, Netherlands. [3]School of Geography, University of Leeds, Leeds, UK. [4]Department of Geograhpy, University College London, London, UK.
Email: julia.jones@bangor.ac.uk

Illegal deforestation takes place within the Serranía de la Macarena National Natural Park in Colombia, one of the countries where carbon offsetting projects are failing to conserve tropical forests.

market for forestry and land-use, with an estimated global value of $1.3 billion USD.

West *et al.* found that most projects did not substantially reduce deforestation, and the few that did, reduced it much less than had been claimed. Furthermore, for a subset of 18 projects for which publicly available information was available, the credits have been used to offset almost three times more carbon dioxide emissions than were avoided by the projects. A further 47 million misleading credits exist and may be sold as "offsets" in the future.

West *et al.* used synthetic control methods (*4*) to robustly estimate how much the REDD+ projects reduced deforestation. They compared this with the predictions made by using the Verified Carbon Standard methods, which is the basis on which the number of credits was sold. Measuring deforestation is relatively straightforward with satellite technology; however, calculating the extent to which REDD+ projects have decreased deforestation is more complicated. It relies on estimating a counterfactual (*5*): how much deforestation would have occurred in the absence of the project.

For each REDD+ project, West *et al.* identified a set of potential control areas of a similar size and characteristics as those of the project area and constructed a synthetic control. This is a weighted average of potential controls, chosen to be as similar as possible compared with the project area in terms of both predictors of deforestation and, crucially, in trends in deforestation before the project started. The projects' outcome is calculated as the difference in deforestation between the synthetic REDD+ project control and the actual REDD+ project, during the years the project was active. As well as calculating the effect of individual projects on deforestation, they also estimate the average effect for all the projects in a country or region and vary how they select potential controls. These extensive robustness checks provide additional confidence in the results.

Why have the methods used by REDD+ projects failed to produce credible estimates of avoided deforestation? The Verified Carbon Standard provides flexibility in how projects estimate a baseline (usually calculated from the extrapolation of historical deforestation) against which deforestation in REDD+ projects is compared and credits are issued. West *et al.* note that this flexibility, together with a tendency to locate projects in areas with low background deforestation (*6*), has resulted in a situation in which projects predict higher deforestation than in practice would have occurred in the project area. This makes REDD+ projects appear more successful at reducing deforestation than they were, and thus more lucrative.

The implications of West *et al.*'s findings are far reaching. Misleading offsets carry negative consequences for the climate because they are not offsetting the emissions released, for forest conservation because they are not reducing deforestation as much as claimed, and for the future finance of forest conservation because the reputational risks of being tainted by accusations of greenwash (misleading or deceptive environmental claims) may deter future investments (*7*). Because of the importance of nature-based solutions to stabilizing the climate—for example, by reducing tropical deforestation (*8*)—the conclusions of West *et al.* have serious consequences for global temperatures.

Some argue that offsets that use avoided emissions are a flawed approach to tackling the climate crisis (*9*) because balancing fossil-fuel emissions with those from land-use change does not achieve the cuts needed to stabilize the climate system (*10*). Furthermore, offsetting reinforces global inequalities because high-income individuals and companies continue to pollute, whereas low-income individuals and countries bear the costs associated with reducing emissions.

Regarding the credibility of the voluntary carbon market, the study of West *et al.* shows that major changes are needed in how credits are calculated. Yet this alone will not protect tropical forests. Deforestation is primarily driven by demand for land to produce agricultural commodities (*11*). This demand can be reduced by reducing food waste from farm to fork, increasing yields on existing agricultural lands, and reducing demand for meat and dairy by high consuming groups, because animal-produced food uses land inefficiently (*12*). Companies would also need to be regulated to ensure deforestation-free commodity supply chains (*13*). Additionally, countries and communities will require income from land that remains forested, rather than being used for agriculture, which will need to come from both private and public sources. With a concerted effort, it should be possible to halt tropical deforestation and its concomitant carbon emissions. Regardless of future developments in policy, independent tests of carbon accounting methods will be needed to evaluate the efficacy of approaches to forest conservation, as West *et al.* clearly demonstrate. ■

REFERENCES AND NOTES

1. P. Friedlingstein *et al.*, *Earth Syst. Sci. Data* **14**, 4811 (2022).
2. Intergovernmental Panel on Climate Change (IPCC), *The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (Cambridge Univ. Press, 2021).
3. T. A. P. West *et al.*, *Science* **381**, 873 (2023).
4. A. Abadie, A. Diamond, J. Hainmueller, *Am. J. Pol. Sci.* **59**, 495 (2015).
5. P. J. Ferraro, *New Dir. Eval.* **122**, 75 (2009).
6. P. Delacote, G. Levelly, G. Simonet, *Resour. Energy Econ.* **67**, 101277 (2022).
7. A. Balmford *et al.*, *Science* **380**, 466 (2023).
8. C. A. J. Girardin *et al.*, *Nature* **593**, 191 (2021).
9. M. Allen *et al.*, *The Oxford Principles for Net Zero Aligned Carbon Offsetting* (Univ. Oxford, 2020). https://www.smithschool.ox.ac.uk/sites/default/files/2022-01/Oxford-Offsetting-Principles-2020.pdf
10. IPCC, *Mitigation of Climate Change. Contribution of Working Group III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (Cambridge Univ. Press, 2022).
11. F. Pendrill *et al.*, *Science* **377**, eabm9267 (2022).
12. J. Poore, T. Nemecek, *Science* **360**, 987 (2018).
13. E. F. Lambin, P. R. Furumo, *Annu. Rev. Environ. Resour.* **48**, annurev-environ-112321-121436 (2023).

10.1126/science.adj6951

NEUROSCIENCE

Good vibrations

The role of a mechanosensitive ion channel in sexual behavior is unveiled

By **Arlene J. George** and **Victoria E. Abraira**

Sexual reproduction is essential for the survival of most species and is the principal driver of animal behavior and evolution. Mammals evolved to have elaborate mating rituals that involve visual, auditory, and olfactory displays. One of the most important elements of mating and copulation is touch, and for many humans, vibration too. Piezo-type mechanosensitive ion channel component 2 (PIEZO2) is a fundamental transducer of touch and vibration for the sensory neurons that innervate the skin. Although there is considerable literature on the role of PIEZO2 in various bodily functions, less is known about its role in genital touch. On page 906 of this issue, Lam *et al.* (*1*) reveal the necessity of PIEZO2 for sexual touch and copulation in mice and humans. Their results both illustrate how sensory neurons in the genital sub-regions can detect gentle stimuli that promote sexual response, and suggest possible therapeutic targets for sexual dysfunction.

Touch is processed through mechanoreceptors, which are sensory neurons that are embedded in the skin (*2*). These sensory neurons comprise high-threshold mechanoreceptors (HTMRs), which are excited by painful touch stimuli, and low-threshold mechanoreceptors (LTMRs), which mediate sensation of innocuous or soft touch. LTMRs transmit touch information through Aβ-, Aδ-, and C-fibers. Aβ- and Aδ-LTMRs are fast-conducting fibers that innervate specialized endings in the skin, including hair follicles, that allow the differentiation of textures, and sense vibration. C-fibers are unmyelinated and slow-conducting, and C-tactile fibers are thought to transmit pleasurable social touch (*3*). The cell bodies of HTMRs and LTMRs reside in the dorsal root ganglia and their axons enter the spinal cord, where touch information is organized, processed, and relayed to the brain. The genitalia are innervated by mechanoreceptors, but they have a higher density of free nerve endings, which are unmyelinated pain receptors that carry touch information through C-fibers (*4*).

Using calibrated filaments of graded stiffness, Lam *et al.* found that perineal or subgenital stimulation predominately activates LTMRs, with ~60% of them responding to gentle stimuli in both male and female mice. By contrast, few HTMRs responded to even the stiffest fibers. To provide a better understanding of the anatomy of the perineal circuit, the authors used viral tracers injected into genital subregions: perineum, prepuce, male glans, and vaginal opening. They found non-overlapping labeling of neurons in sacral dorsal root ganglia, indicating that each subregion was innervated by a distinct group of neurons. These tracing studies also showed that neurons innervating the different subregions projected to distinct parts of the dorsal horn of the spinal cord. Furthermore, the authors showed that the perineum is innervated by LTMRs that have specialized endings wrapped around hair follicles.

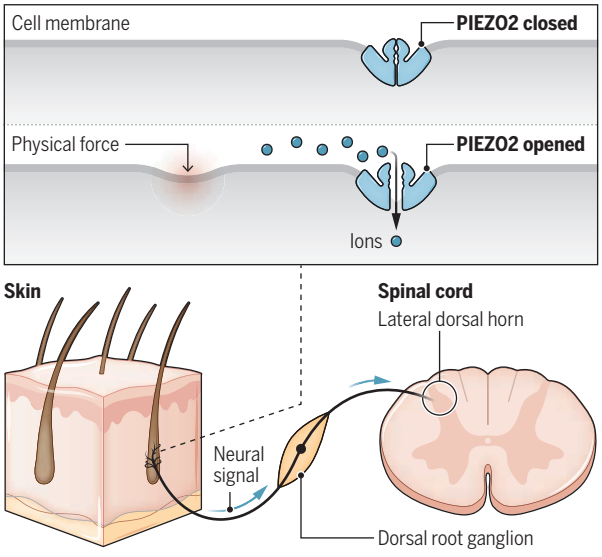
PIEZO1 and PIEZO2 are mechanosensitive ion channels that are enriched in somatosensory neurons and evolutionarily conserved across invertebrates and vertebrates (*5*). PIEZO2 is usually in a closed state and opens in response to mechanical stimulation. It has a role in light-touch sensation by sensory neurons and is expressed within mechanoreceptors such as Merkel's discs where it translates the mechanical stretch of skin into electrical impulses sent to the brain. Lam *et al.* parsed out PIEZO2's function using a mouse model with *Piezo2* deletion from the neck down (*Piezo2*^{Hoxb8} mice). The withdrawal response to perineal mechanical stimulation was lost in these mice, even with the highest force tested; the touch-induced erection reflex was also impaired. However, the response of Piezo2^{Hoxb8} mice to painful pin-prick stimuli to the perineum was not different from that of normal littermates. Thus, PIEZO2 is important for triggering behavioral responses to gentle perineal touch in mice (see the figure).

To understand the type of LTMR fibers involved, the authors also generated a mouse model with *Piezo2* removed from pain receptors only (*Piezo2*^{Scn10a} mice) and found that they had large erection reflex deficits similar to those of Piezo2^{Hoxb8} mice. Sodium channel protein type 10 subunit α (*Scn10a*)—a marker for pain receptors—and PIEZO2 were co-expressed with a C-tactile fiber marker in the sacral ganglia, which indicates that perineal C-tactile fibers regulate the erection reflex. *Piezo2*^{Hoxb8} and *Piezo2*^{Scn10a} males were not able to successfully mate despite displaying normal sexual motivation. *Piezo2*^{Hoxb8} females did not display proper receptive behavior, preventing successful mating and indicating that PIEZO2-dependent input is essential for mating in males and females.

Sexual function and reproduction are dependent on the integrity of the hypothalamic-pituitary-gonadal axis, which controls reproductive behavior through the secretion of hormones by tightly regulated neural networks, resulting in the production and release of follicle-stimulating hormone, testosterone, and estrogens from the gonads (*6*). In particular, oxytocinergic neurons in the hypothalamus project to other areas of the

Detecting sexual touch

Touch in the genital area mechanically opens piezo-type mechanosensitive ion channel component 2 (PIEZO2) receptors in lanceolate and circumferential nerve endings surrounding hair follicles, which allows ions to flood into the neuron. This influx of ions causes an electrical pulse to be sent through low-threshold mechanoreceptors to the dorsal horn of the spinal cord. This signal induces a pleasurable sensation when it reaches the brains of mice.



Department of Cell Biology and Neuroscience, Rutgers University, New Brunswick, NJ, USA.
Email: victoria.abraira@rutgers.edu

brain and the spinal cord and are necessary for penile erection and copulatory behavior (*7*). Many years of research have suggested that hormones, specifically testosterone, are the primary driver of copulatory success. However, Lam *et al.* did not find differences in testosterone or follicle-stimulating hormone expression or changes in female estrous cyclicity in the *Piezo2* mutant mice, indicating that PIEZO2-based touch activation for sexual behavior does not involve hypothalamic activation or sex hormones.

Because *Piezo2* is conserved across vertebrates, understanding its function in sexual behavior could have clinical relevance. Lam *et al.* performed clinical evaluations in individuals with PIEZO2-deficiency syndrome, which is a rare condition that results from mutations in the *PIEZO2* gene. PIEZO2 is expressed in many different types of tissues and therefore mutations in the gene can cause an array of symptoms including body position and joint dysfunction, scoliosis, hip dysplasia, and muscular atrophy with perinatal respiratory distress (*8*). Lam *et al.* found that individuals with loss of PIEZO2 had a lack of proprioception, absent vibration sensing, elevated touch threshold, and scoliosis but normal puberty (an indication of normal hypothalamic function), which is comparable with the clinical literature (*8*). Individuals with PIEZO2-deficiency were sexually active and able to become aroused but did not respond to gentle genital stimulation, which is consistent with the *Piezo2*^{Hoxb8} mouse model. Although there are no pharmacological treatments to compensate for PIEZO2 deficiency, understanding the function of this ion channel could inform future approaches to treating sexual dysfunction.

Other avenues for future work include determining whether the PIEZO2 sensory pathway is sufficient for sexual function and which sensory neuron subtypes are involved in this spinal circuit. It is also not clear which sensory endings in the genitalia sense vibrational touch, which was absent in individuals with PIEZO2 deficiency syndrome. Future studies in mice are needed to parse out the PIEZO2 circuit of sexual touch and how peripheral perigenital sensation is represented in the spinal cord and brain. ■

GRAPHIC: A. FISHER/*SCIENCE*

REFERENCES AND NOTES

1. R. M. Lam *et al.*, *Science* **381**, 906 (2023).
2. W. Olson, P. Dong, M. Fleming, W. Luo, *Wiley Interdiscip. Rev. Dev. Biol.* **5**, 389 (2016).
3. V. E. Abraira, D. D. Ginty, *Neuron* **79**, 618 (2013).
4. R. D. Johnson, Z. Halata, *J. Comp. Neurol.* **312**, 299 (1991).
5. A. T. Chesler *et al.*, *N. Engl. J. Med.* **375**, 1355 (2016).
6. G. W. Harris, *BMJ* **2**, 559 (1951).
7. A. Argiolas, M. R. Melis, *Prog. Neurobiol.* **76**, 1 (2005).
8. S. L. Alper, in *Current Topics in Membranes*, P. A. Gottlieb, Ed. (Academic Press, ed. 79, 2017), pp. 97–134.

10.1126/science.adj8674

ASTRONOMY

Thin jets underlie the solar wind

Solar Orbiter images reveal widespread magnetic plasma jets at the roots of the solar wind

By **Ignacio Ugarte-Urra** *and* **Yi-Ming Wang**

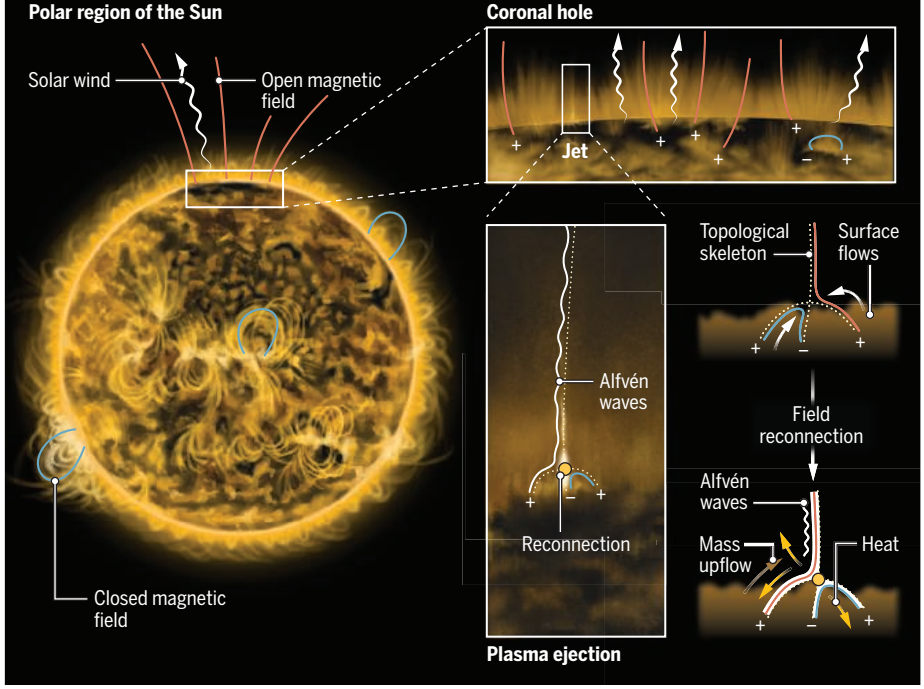
The Solar Orbiter spacecraft was launched in 2020 by the European Space Agency to study how solar activity drives changes in the inner part of the heliosphere, the region that extends from the Sun to the surrounding interstellar medium. The Solar Orbiter is fitted with a suite of in situ and remote-sensing instruments that allow measurement of the properties of the particles and magnetic field traversed by the spacecraft while simultaneously observing the magnetic and radiative sources at the Sun. On page 867 of this issue, Chitta *et al.* (*1*) present observations by the Extreme Ultraviolet Imager (EUI) instrument that identify widespread collimated outflows (jets) at the base of the solar wind that suggest the ubiquitous small-scale dissipation of magnetic energy. The solar wind, a continuous stream of charged particles flowing outward from the Sun, is an important component of space weather that influences Earth's magnetic environment.

Space Science Division, Naval Research Laboratory, Washington, DC, USA. Email: ignacio.ugarte-urra@nrl.navy.mil; yi.wang@nrl.navy.mil

Solar wind was described in the 1950s as "solar corpuscular radiation" [(*2*), p. 274] and identified through its interaction with the ion tails of comets. In 1958, astrophysicist Eugene Parker proposed that the thermal pressure gradients associated with the Sun's million-degree outer atmosphere or "corona" would cause it to continually expand outward, resulting in the solar wind (*3*). This prediction was subsequently confirmed by NASA's Mariner 2 mission (*4*). The exact processes involved in generating and accelerating the solar wind have been debated ever since. The proposed mechanisms have generally fallen into two catego-

Jets at the Sun's poles

The polar observations by the Extreme Ultraviolet Imager instrument on the Solar Orbiter spacecraft reveal plasma ejections (jets) that permeate coronal holes (areas of open magnetic fields). Jets arise from regions of magnetic reconnection. This reconfiguration of the magnetic fields also generates heat, Alfvén waves, and mass upflows that, together with the jets, provide energy and momentum to the solar wind.



ries (*5*): those that involve magnetic field reconnection, a reconfiguration of the Sun's field that releases its built-up stresses; and those that involve the dissipation of energy from magnetohydrodynamic waves, which are pervasive oscillations in the coronal magnetic field.

The EUI takes advantage of the highly elliptical orbit of the Solar Orbiter and a perihelion approach of 0.28 astronomical units (Earth-Sun distance) to record some of the highest-resolution images of the corona to date. Chitta *et al.*'s analysis of 200-km-resolution observations of the Sun's south pole has unveiled faint widespread and frequent jets of coronal plasma that may contribute substantially to the solar wind energy flux. These narrow plasma ejections, a few hundred kilometers across, originate from sites where magnetic reconnection is observed (*6*). They are a signature of "interchange" reconnection between the open magnetic field lines, which extend outward into the heliosphere, and small closed fields that connect back to the surface (see the figure). This rearrangement of the magnetic field, which is driven by convective motions at the surface, converts magnetic energy into plasma outflows, magnetohydrodynamic waves, and heat.

With the steadily improving sensitivity and spatial resolution of extreme ultraviolet (EUV) instruments such as the Atmospheric Imaging Assembly (AIA) on the Solar Dynamics Observatory (*7*), the Solar Ultraviolet Imager (SUVI) on the Geostationary Operational Environmental Satellite R series (GOES-R) (*6*), and now EUI on the Solar Orbiter, it is clear that coronal holes, or areas of open magnetic fields, are permeated by these narrow jets, which have temperatures close to 1 MK. This does not necessarily mean that the solar wind is simply a collection of jets; the underlying magnetic reconnection events may also heat the top of the dense chromosphere below the corona and evaporate material to supply the bulk of the solar wind mass flux.

The reconnection and magnetohydrodynamic wave scenarios for solar wind energization are not mutually exclusive. The small-scale reconnection events that produce the observed jets will also generate Alfvén waves, which are transverse oscillations of the magnetic field. The solar wind from the interiors of coronal holes reaches speeds of 600 to 800 km/s, exceeding the speed of a typical jet (100 km/s), which is also much less than the escape velocity (the speed required to overcome a body's gravitational pull) at the Sun's surface (620 km/s). To achieve high speeds, energy must be deposited up to and beyond the point where the wind velocities become supersonic, which is typically at distances of 2 to 3 solar radii from the Sun's center. The outward-propagating Alfvén waves provide a way of transmitting energy to the outer corona. Numerical models show that if some of the waves are reflected, interactions between incoming and outgoing waves will result in the deposition of energy and momentum into the solar wind (*4*).

At heliocentric distances beyond 10 solar radii, NASA's Parker Solar Probe (launched in 2018) measured velocity spikes and temporary reversals in the magnetic field (*8*), called switchbacks. It has been proposed that these small-scale field reversals are injected at the underlying reconnection sites (*9*) or, alternatively, that they develop as the Alfvén waves propagate outward and evolve into a turbulent state (*10*). The jetting activity described by Chitta *et al.* shows that interchange reconnection is heating the base of the corona and helping to drive the solar wind mass flux. A remaining puzzle is that these reconnection events require small closed magnetic fields and thus the presence of small-scale fields of opposite polarity to that of the open field lines. Magnetic measurements show very little opposite-polarity flux at the feet of the open field lines (*11*). This "missing" flux may eventually be detected using magnetographs with improved sensitivity, such as those now coming into operation at Big Bear Solar Observatory in Big Bear, California, and the Daniel K. Inouye Solar Telescope in Maui, Hawaii, the world's largest solar telescope.

Fast-cadence, high-resolution spectroscopy with comprehensive temperature coverage will also be key to understanding the fundamental heating processes and the transfer of mass from the chromosphere into the corona. Moreover, it will be important to resolve the accompanying wave motions and to determine whether the elemental composition at the base matches the in situ measurements in the solar wind. The upcoming Japan Aerospace Exploration Agency and NASA spectroscopic missions, Solar-C EUV High-throughput Spectroscopic Telescope (EUVST) and Multi-Slit Solar Explorer (MUSE), respectively, promise to provide a definitive battery of diagnostics to peer into the roots of the solar wind. ■

REFERENCES AND NOTES

1. L. P. Chitta *et al.*, *Science* **381**, 867 (2023).
2. L. Biermann, *Z. Astrophys.* **29**, 274 (1951).
3. E. N. Parker, *Astrophys. J.* **128**, 664 (1958).
4. T. I. Gombosi *et al*, *Living Rev. Sol. Phys.* **15**, 4 (2018).
5. S. R. Cranmer, A. R. Winebarger, *Annu. Rev. Astron. Astrophys.* **57**, 157 (2019).
6. N.-E. Raouafi *et al.*, *Astrophys. J.* **945**, 28 (2023).
7. H. Tian *et al.*, *Astrophys. J.* **736**, 130 (2011).
8. D. Bale *et al.*, *Astrophys. J.* **923**, 174 (2021).
9. J. F. Drake *et al.*, *Astron. Astrophys.* **650**, A2 (2021).
10. J. Squire *et al.*, *Astrophys. J. Lett.* **891**, L2 (2020).
11. Y.-M. Wang, *Astrophys. J.* **904**, 199 (2020).

10.1126/science.adj8002

CELL BIOLOGY

Catching proteins for degradation

A ubiquitin-independent pathway targets nuclear proteins to the proteasome

By **Charlotte M. Schilling** *and* **Eilika Weber-Ban**

The breakdown of specific proteins at precise times affords control over crucial cellular processes. The molecular machine carrying out this degradation in eukaryotes is the 26*S* proteasome (*1*). Ubiquitin ligases provide selectivity by attaching the small protein ubiquitin to specific proteins, thus marking them for destruction by the 26*S* proteasome (*1*). Ubiquitin-independent recruitment to the 26*S* proteasome exists, but the processes involved are less understood and comparatively understudied. On page 849 of this issue, Gu *et al.* (*2*) report a ubiquitin-independent mechanism for targeting important nuclear proteins for proteasomal degradation. At the center of this pathway is a recruiter protein called midnolin that captures substrates with its Catch domain. The discovery suggests that ubiquitin-independent proteasomal degradation could be the preferred degradation route for a specific set of substrates.

Immediate early genes (IEGs) are rapidly and transiently expressed in many cells in response to extracellular stimuli (*3*). The proteins encoded by IEGs are localized to the nucleus because their primary function is to regulate the expression of other genes. The transient presence of IEG proteins is crucial for processes such as synaptic plasticity in neurons or the responses of immune cells; however, the mechanisms controlling their degradation are poorly understood.

Gu *et al.* used a protein stability reporter system to assess the stability against degradation of two well-characterized IEG proteins in cultured human cells. They carried out a genome-wide loss-of-function screen to identify genes that affect the amounts of the two reporter IEG proteins. The gene that, when disrupted, led to the strongest increase in their levels, encodes the nuclear protein midnolin (*4*). Subsequent experiments con-

Institute of Molecular Biology and Biophysics, ETH Zurich, Zurich, Switzerland. Email: eilika@mol.biol.ethz.ch

firmed that IEG proteins accumulated in cells (including neural and immune cells) mutated to be deficient in midnolin, whereas overexpression of midnolin led to IEG protein depletion. Treatment of midnolin-overexpressing cells with a proteasome inhibitor also led to accumulation of IEG proteins, but inhibition of ubiquitination did not have this effect. IEG protein degradation thus depended on both midnolin and the proteasome but not on ubiquitination.

These findings are consistent with a report that the fruit fly (*Drosophila melanogaster*) ortholog of midnolin, Stuxnet, binds the subunit Polycomb of the Polycomb repressive complex 1 (an epigenetic repressor) (*5*) and mediates the ubiquitin-independent proteasomal degradation of Polycomb (*6*). Another example of a specialized recruitment factor is antizyme 1, which triggers ubiquitin-independent proteasomal degradation of ornithine decarboxylase—the first enzyme in polyamine biosynthesis (*7*). By contrast, Gu *et al.* showed that midnolin interacts with numerous protein clients through a common recruitment mechanism. A large-scale stability screen assessed the levels of proteins encoded by nearly 12,000 genes in midnolin-deficient and midnolin-overexpressing human cell lines, of which 508, mostly DNA-binding proteins or transcription factors, were stabilized in the absence of midnolin.

The midnolin structure generated by the artificial intelligence (AI)–based structure-prediction program AlphaFold, comprises three stable domains: a ubiquitin-like domain located toward the N terminus, a two-part pseudosymmetric Catch domain, and a long C-terminal α helix (αHelixC). Although all three domains were required for IEG protein degradation, coimmunoprecipitation experiments revealed that client binding occurs through the Catch domain, and association with the 26*S* proteasome depends on the αHelixC but not on the ubiquitin-like domain. The latter finding is unexpected because shuttle factors—such as Rad23 and ubiquilins—that help to deliver certain ubiquitinated substrates to the proteasome use ubiquitin-like domains for binding (*1, 8*). Furthermore, Stuxnet could no longer interact with the proteasome when the ubiquitin-like domain was deleted in *D. melanogaster* cells, suggesting a role for this domain in binding to the proteasome (*6*).

Notably, proteins containing ubiquitin-like domains not only bind 26*S* proteasomes but also stimulate their activity (*9*). Therefore, it is possible that both the αHelixC and the ubiquitin-like domain of midnolin interact with the proteasome, the ubiquitin-like domain stimulating proteasome activity by docking to one of the ubiquitin receptor subunits and the αHelixC ensuring stable complex formation. Interaction of midnolin with the 26*S* proteasome through αHelixC might involve the coiled-coil domains on the adenosine triphosphate (ATP)–hydrolyzing subcomplex of the 26*S* proteasome. In bacteria, the proteasome recruits substrates through these coiled-coil domains and this involves the formation of a shared three-stranded coiled-coil with a long helix on the factor targeting substrates for degradation (*10*). However, a molecular understanding of midnolin-mediated recruitment to the proteasome awaits further mechanistic and structural analysis.

Gu *et al.* offer an intriguing explanation of how the Catch domain of midnolin recruits its client proteins. The two subdomains each contribute a β hairpin to form an incomplete β sheet. A groove between the two hairpins serves as the acceptor site for an additional strand donated by client proteins carrying a recognition motif for degradation (degron) with a propensity to form β strands (see the figure). This hypothesis was derived from an AlphaFold structure model of an IEG protein bound to midnolin. Among the 508 proteins that were stabilized in midnolin-deficient cells, 205 showed predictions of β-sheet complementation in AlphaFold models. Mutations disrupting or deleting the β-strand degrons in several client proteins led to their accumulation in human cells.

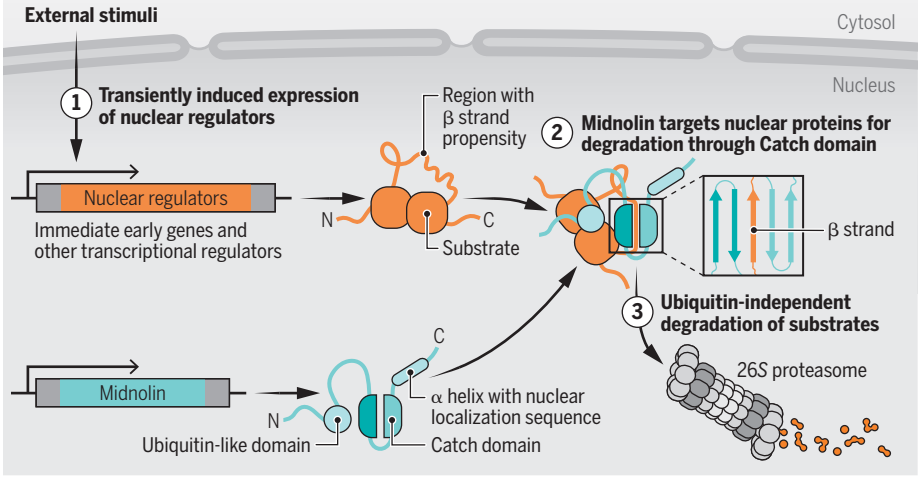
Binding mechanisms involving β-strand additions are well suited for interactions with a large number of proteins because they involve both nonspecific main-chain interactions with neighboring strands and specific interactions via side chain residues (*11*). Furthermore, β-strand degrons do not require three-dimensional structural context for binding, so proteins exhibiting regions of disorder—characteristics of many degradation substrates—can be recruited by this mechanism.

A substantial portion of midnolin is predicted to be intrinsically disordered. These regions might adopt structure upon interaction with the proteasome or with certain client proteins. However, proteasomal degradation substrates must feature an unstructured region to initiate substrate processing (*12*), so it is possible that midnolin is degraded along with its cargo and that the disordered regions facilitate this process.

Ubiquitin-independent proteasomal degradation has long been considered the exception to the rule. This notion is challenged by the discovery of common mechanistic principles for the ubiquitin-independent recruitment of many nuclear proteins. Future work should include in vitro biochemical characterization of the process and structural analysis of the participating machinery. ■

Midnolin-dependent degradation of nuclear regulators

In response to external stimuli, expression of immediate early gene proteins and other transcriptional regulators is transiently induced (**1**). Midnolin targets many of these proteins for degradation by binding to regions with β strand propensity through its Catch domain (**2**). Subsequent ubiquitin-independent degradation of substrates by the 26*S* proteasome, possibly together with midnolin, ensures tight regulation of protein levels (**3**).



GRAPHIC: N. BURGESS/*SCIENCE*

REFERENCES AND NOTES

1. X. Chen *et al.*, *FEBS J.* **288**, 5231 (2021).
2. X. Gu *et al.*, *Science* **381**, eadh5021 (2023).
3. B. Pérez-Cadahía *et al.*, *Biochem. Cell Biol.* **89**, 61 (2011).
4. M. Tsukahara *et al.*, *Gene* **254**, 45 (2000).
5. U. Grossniklaus, R. Paro, *Cold Spring Harb. Perspect. Biol.* **6**, a019331 (2014).
6. J. Du *et al.*, *Dev. Cell* **37**, 507 (2016).
7. C. Kahana, *J. Biol. Chem.* **293**, 18730 (2018).
8. B. C. Lin *et al.*, *FEBS J.* **289**, 6132 (2022).
9. G. A. Collins, A. L. Goldberg, *Proc. Natl. Acad. Sci. U.S.A.* **117**, 4664 (2020).
10. T. Wang *et al.*, *Nat. Struct. Mol. Biol.* **17**, 1352 (2010).
11. H. Remaut, G. Waksman, *Trends Biochem. Sci.* **31**, 436 (2006).
12. T. Tomita, A. Matouschek, *Protein Sci.* **28**, 1222 (2019).

10.1126/science.adj8230

RETROSPECTIVE

John B. Goodenough (1922–2023)

A giant in the fields of solid-state chemistry and physics

By **Clare P. Grey**[1] *and* **Laura H. Lewis**[2]

John Bannister Goodenough, groundbreaking materials scientist, died on 25 June at the age of 100. John made seminal contributions to solid-state chemistry, physics, and engineering. In 2019, he shared the Nobel Prize in Chemistry for the development of lithium-ion batteries. His revolutionary insights into the fundamental physical properties of materials helped enable the wireless and artificial intelligence revolutions and advanced the science needed to help reduce carbon emissions.

A New Englander in both upbringing and demeanor, John was born on 25 July 1922 in Jena, Germany, and raised in Connecticut. He received a BA in mathematics from Yale University in 1943 and served as a World War II US Army Air Forces captain from 1942 to 1948. In 1952, he earned his PhD in physics from the University of Chicago for—in his words—a "modest thesis," written under the guidance of theoretical physicist Clarence Zener, that investigated structural changes in alloys as a function of electronic configuration. Eschewing a traditional academic career, John then joined the Massachusetts Institute of Technology's Lincoln Laboratory. He moved to become head of the Inorganic Chemistry Laboratory at the University of Oxford in 1976 and then 10 years later to the University of Texas at Austin, where he remained a chaired professor in the mechanical engineering department until his death.

John's early understanding of solids from a theoretical perspective served as the foundation for his later experimental inquiries, allowing him to view existing materials in new ways and providing a blueprint for creating new types of materials with designer functionalities. He began his influential work at the height of the Cold War. Lincoln Laboratory, established by the Air Force in 1950, had developed the first US air-defense system, combining radar, communications, and the early digital computer to detect and track hostile aircraft. In pursuit of faster computer memory, John studied magnetic oxides, providing insight extending from chemical bonding and spin orientation all the way to microstructure and performance. In 1955, he published a series of predictions describing how two transition-metal magnetic cations interact through the intervening oxygen anions. These rules, updated in 1959 by Junjiro Kanamori and now known as the Goodenough-Kanamori rules, helped shape our understanding of magnetic couplings in magnetic systems and are still used today.

Building on his foundational ideas, John explored phenomena such as localized versus band descriptions of electronic transport as well as Jahn-Teller distortions and spin-orbit coupling effects in oxides. These efforts ultimately led to the development of random-access memory (RAM), which remains a fundamental component of computing. His work also sparked ideas that led to breakthroughs in superconductors, rechargeable batteries, and heterogeneous catalysis.

In the mid-1970s, motivated by the energy crisis, John turned his attention to methods for capturing and storing energy. In 1975, he published the first report on the very high sodium-ion conductivity within a family of sodium zirconium phosphates. He identified the need for wide-open windows to allow rapid hopping between connected sites ("bottlenecks"). This work unlocked the field of sodium (Na) superionic conductors (NASICONs), a range of Na- and lithium (Li)–ion conductors that have now been widely studied both as electrode materials and ionic conductors for liquid-based and solid-state batteries, respectively. It also led him to recognize that covalency effects within the skeleton structures alter the redox potentials of the redox-active ions. His early ideas on this phenomenon, which he coined the "inductive effect," are still widely applied in the field. John demonstrated reversible extraction of Li from layered oxides at high voltage. He noted that "cobalt is expensive, with little attraction for the manufacturer." Ironically, 10 years later, Sony used lithium cobalt oxide to make the first reversible Li-ion battery. In 1983, Michael Thackeray arrived from South Africa with a bag of iron-based spinel gems. Together, he and John developed manganese spinels, not as magnetic structures or hopping semiconductors but as high-rate Li-ion cathode materials for Li-ion batteries. In 2019, John was recognized for his contributions to the development of Li-ion batteries with the Nobel Prize.

High-temperature copper-oxide superconductors burst onto the scene in 1986, and John and his then-postdoc Arumugam Manthiram at the University of Texas began their long collaboration by studying synthesis and doping in yttrium barium copper oxide. In 1997, John and postdoc Akshaya K. Padhi developed the olivine lithium iron phosphate as a positive electrode material, which is now used commercially in cheap and high-rate batteries. John continued to advance our understanding of magnetic and orbital ordering coupling in perovskite materials with colossal magnetoresistance, of enduring interest for spintronics, for the rest of his life.

C.P.G. met John as a student and recalls him bravely teaching second-year undergraduates radically new concepts, which remain relevant to her work today. She later collaborated with John through the US Department of Energy battery research program. L.H.L. was a PhD student studying under John's guidance in the mechanical engineering department at the University of Texas. In his button-down pink shirt and trademark bow tie, he provided formal but kind and patient supervision that conveyed his very high expectations. John required his research group students to attend all of his graduate courses, a time-consuming obligation that, in retrospect, was a great gift and privilege. It was during those textbook-free classes that he communicated his unique, intuitive understanding of how nature operates at the atomic scale. Interactions with John were marked by his infectious laugh. Slow, loud, and a bit braying, this laugh could convey surprise, delight, incredulity, bemusement, or irony and provided a useful glimpse of his thoughts to those who knew him well enough to interpret it.

John's work was both interdisciplinary and international—he nurtured and collaborated with scientists all over the world. His contributions to science spanned 71 years, and his legacy will be even more long-lived. The technologies he made possible will define the 21st century and beyond. ■

10.1126/science.adj9263

[1]Yusuf Hamied Department of Chemistry, University of Cambridge, Cambridge, UK. [2]College of Engineering, Northeastern University, Boston, MA, USA.
Email: cpg27@cam.ac.uk; lhlewis@northeastern.edu

POLICY FORUM

CLIMATE POLICY

Mandatory disclosure would reveal corporate carbon damages

Accurate reporting is critical for markets and climate policies

By **Michael Greenstone[1,2], Christian Leuz[2,3,4], Patricia Breuer[5]**

The US Securities and Exchange Commission recently proposed a rule that would mandate that public companies report their greenhouse gas (GHG) emissions. This follows similar efforts in the European Union (EU) and United Kingdom. One rationale is that disclosure will provide information on material risks to investors, making it evident which firms are most exposed to future climate policies. In addition, some believe that reporting will galvanize pressure from companies' key stakeholders (e.g., customers and employees), leading them to voluntarily reduce their emissions. This reasoning is in line with evidence for financial markets (*1*) and disclosure mandates that form the third wave of environmental policy, which follows a wave of direct regulation and a wave of market-based approaches (*2*). But what might such disclosure reveal? We provide a first-cut preview of what we might learn about the climate damages caused by each company's GHG emissions by drawing on one of the largest global datasets, which covers roughly 15,000 public companies.

Here, we introduce "corporate carbon damages" as a measure of the total costs to society associated with corporate emissions. For each firm, it is calculated as the product of their carbon dioxide (CO_2)–equivalent direct emissions and the social cost of carbon (SCC)—the monetary value of the damages associated with the release of an additional metric ton of CO_2. To account for differences in firm size and to facilitate across-firm comparisons, we then divide this product by the respective firm's operating profit or sales. With existing datasets, it is not possible to determine who bears the costs or to divide responsibility for these damages between firms and consumers (*3*). We nevertheless refer to them as corporate carbon damages because the emissions come from firm activities. The core finding is that average corporate carbon damages are large, but they vary greatly across firms within an industry, across industries, and across countries.

We argue that widespread mandatory disclosure is critical for confronting the climate challenge for several reasons. Perhaps most importantly, reliable measurement of GHG emissions is the foundation of any meaningful policy to restrict emissions. Additionally, knowledge of the heterogeneity in corporate carbon damages is critical to tackling the distributional and related political economy considerations that frequently cause climate policies to founder. Finally, making heterogeneity in corporate emissions transparent can facilitate across-firm benchmarking by various stakeholders, which could bec an important force that drives continued reductions in corporate emissions.

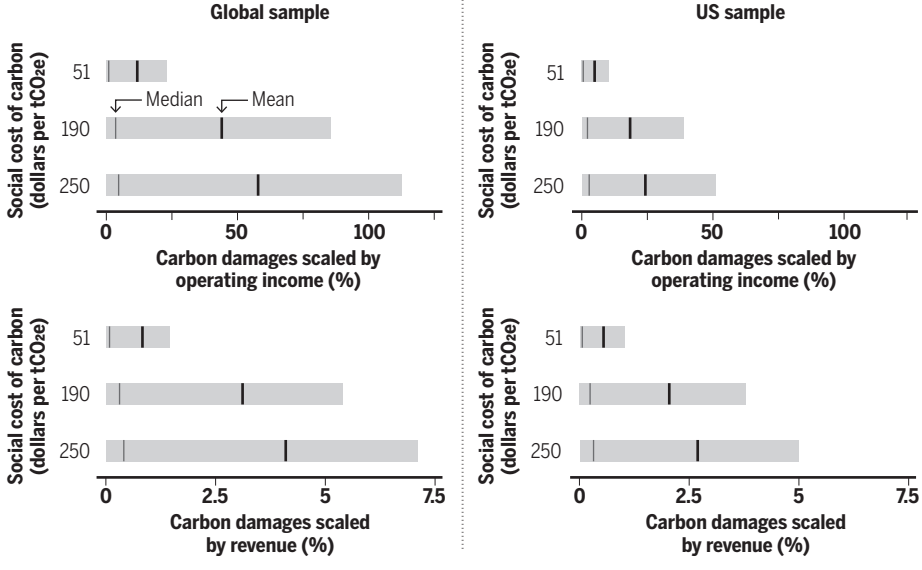
DATA AND METHODS

Our analysis is based on a global sample of reported and estimated corporate GHG emissions provided by Trucost [see supplementary materials (SM) and (*4*)]. We focus on data from 2019 because it is the most recent year that was unaffected by the COVID-19 pandemic; however, we also report results for 2021 that are qualitatively similar (table S2). Our focus is on scope 1 emissions, which are the direct emissions from sources that are owned or controlled by the respective company. This focus avoids issues of double counting that could otherwise arise [e.g., when using indirect emissions from purchased electricity (scope 2) or upstream production inputs and downstream use (scope 3)]. In principle, scope 1 emissions of all companies should add up to the corporate sector's total emissions. Moreover, if all firms globally reported their scope 1 emissions, all corporate emissions would be accounted for, including those from firms that "outsource" to other countries; indeed, outsourcing across country borders provides another rationale for conducting our analysis at the global level.

A weakness of an exclusive focus on scope 1 emissions is that it involves some arbitrariness in assigning carbon damages to firms, industries, and countries. For example, a steel producer that burns fossil fuels on-site would be rated as producing

Corporate carbon damages based on operating income and revenue

Gray horizontal bars indicate ranges from the 10th to 90th percentiles. When computing carbon damages scaled by operating income, only firms for which operating income is positive are used. Observations of the scaled carbon damages are truncated if they are below the first or above the 99th percentile.

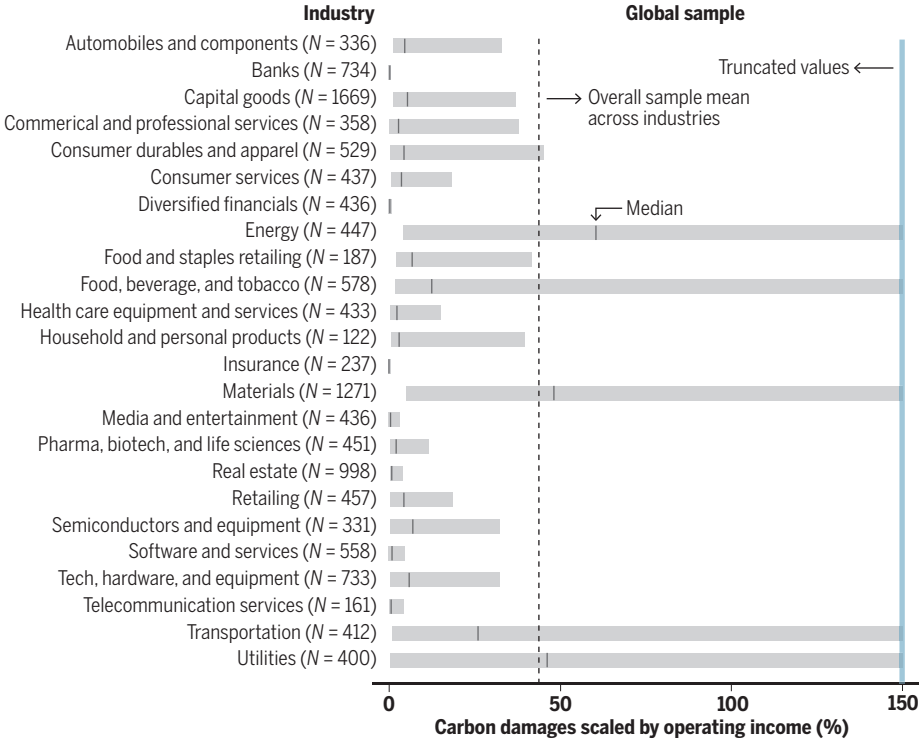


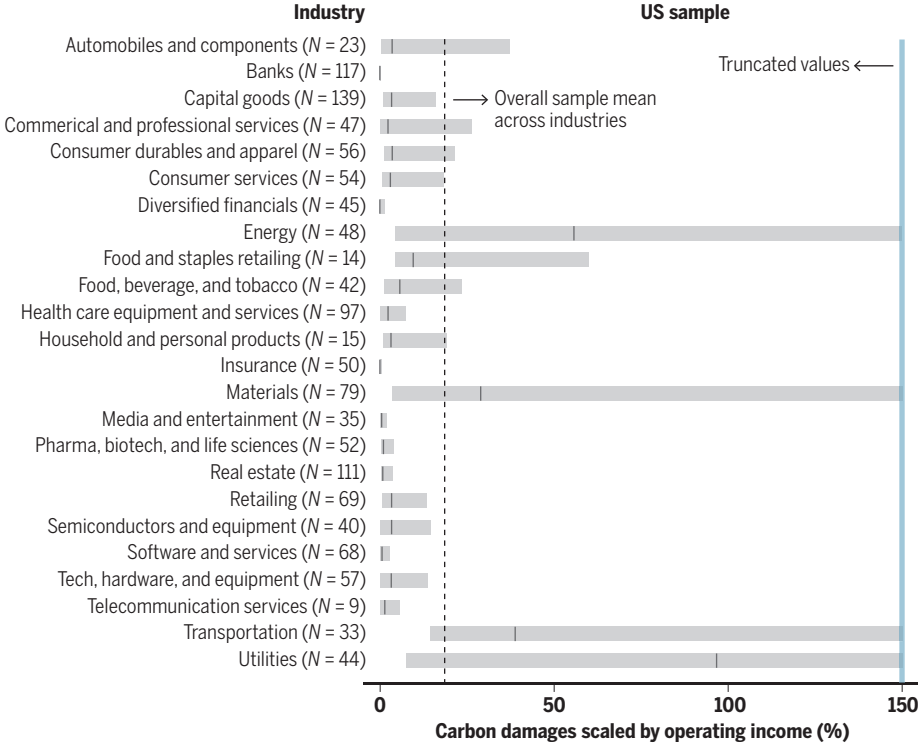
[1]Department of Economics, University of Chicago, Chicago, IL, USA. [2]National Bureau of Economic Research, Cambridge, MA, USA. [3]Booth School of Business, University of Chicago, Chicago, IL, USA. [4]Centre for Economic Policy Research, London, UK. [5]Business School, University of Mannheim, Mannheim, Germany. Email: christian.leuz@chicagobooth.edu

GRAPHIC: D. AN-PHAM/*SCIENCE*

Corporate carbon damages by industry

Carbon damages are shown for a social cost of carbon of $190 per ton of CO_2-equivalent emissions. Gray horizontal bars indicate ranges from the 10th to 90th percentiles. The number of firms (*N*) in each industry group are in parentheses. Graphs are truncated at 150%. In the global (US) sample, the values of the 90th percentile for energy; food, beverage, and tobacco; materials; transportation; and utilities are 383% (234%); 160%; 567% (178%); 358% (201%); and 675% (342%), respectively.


Industry
Global sample
Automobiles and components (N = 336)
Banks (N = 734)
Capital goods (N = 1669)
Commerical and professional services (N = 358)
Consumer durables and apparel (N = 529)
Consumer services (N = 437)
Diversified financials (N = 436)
Energy (N = 447)
Food and staples retailing (N = 187)
Food, beverage, and tobacco (N = 578)
Health care equipment and services (N = 433)
Household and personal products (N = 122)
Insurance (N = 237)
Materials (N = 1271)
Media and entertainment (N = 436)
Pharma, biotech, and life sciences (N = 451)
Real estate (N = 998)
Retailing (N = 457)
Semiconductors and equipment (N = 331)
Software and services (N = 558)
Tech, hardware, and equipment (N = 733)
Telecommunication services (N = 161)
Transportation (N = 412)
Utilities (N = 400)
Truncated values
Overall sample mean across industries
Median
0
50
100
150
Carbon damages scaled by operating income (%)


Industry
US sample
Automobiles and components (N = 23)
Banks (N = 117)
Capital goods (N = 139)
Commerical and professional services (N = 47)
Consumer durables and apparel (N = 56)
Consumer services (N = 54)
Diversified financials (N = 45)
Energy (N = 48)
Food and staples retailing (N = 14)
Food, beverage, and tobacco (N = 42)
Health care equipment and services (N = 97)
Household and personal products (N = 15)
Insurance (N = 50)
Materials (N = 79)
Media and entertainment (N = 35)
Pharma, biotech, and life sciences (N = 52)
Real estate (N = 111)
Retailing (N = 69)
Semiconductors and equipment (N = 40)
Software and services (N = 68)
Tech, hardware, and equipment (N = 57)
Telecommunication services (N = 9)
Transportation (N = 33)
Utilities (N = 44)
Truncated values
Overall sample mean across industries
0
50
100
150
Carbon damages scaled by operating income (%)

much higher damages than one that draws electricity from the grid with the same total associated CO_2 emissions. Similarly, a firm (or country) can reduce its scope 1 emissions by "outsourcing" the emissions-intensive part of its business (or economy) by spinning it off into a separate company (or by regulating emissions domestically so that emissions move to another country). Acknowledging these weaknesses, we use scope 1 emissions to avoid double counting, which is critical if the goal is to compute firm-level carbon damages. We show that the results of our analysis are qualitatively unchanged when we account for differences in power supply and use the sum of scope 1 and 2 emissions instead (tables S9 to S11).

Our study provides and analyzes estimates of corporate carbon damages for 14,879 publicly traded firms across the globe. Our sample accounts for more than 80% of global market capitalization of public companies. We calculate each company's carbon damages as the product of its tons of scope 1 emissions and the SCC. We use three different estimates for the SCC to account for the uncertainties in estimating climate damages. Most of our analyses use $190 per ton of CO_2-equivalent emissions (tCO_2e), which was introduced by the US Environmental Protection Agency in November 2022, but we also report results for a lower value ($51 per tCO_2e) that was used by the Obama administration and a higher value ($250 per tCO_2e) that reflects that the $190 per tCO_2e value does not include areas (e.g., migration and conflict) for which it is generally expected that there will be meaningful climate damages (*5*).

We then normalize the corporate carbon damages to provide a sense of their magnitude at the firm level. One way to scale it is to use output or revenue, akin to what economists do when they compute the labor and capital shares as production inputs. Another way to normalize carbon damages is to express them as a fraction of firms' operating profits, which are, on average, about 17% of firms' revenues; this expresses carbon damages relative to a measure of the value firms create with their products and services for their shareholders.

The idea of monetizing firms' environmental impacts has been around for at least a decade. We can broadly group approaches into three categories: (i) conceptual proposals that extend the income statement by line items reflecting the monetized value of corporate impacts; (ii) studies that suggest or analyze accounting methods for GHG emissions; and (iii) empirical analyses of monetized corporate environmental impacts or carbon taxes. Our analysis of carbon damages differs from related ap-

proaches because of its large global sample of firms from many industries, the values for SCC, and its conceptual underpinnings, such as our focus on scope 1 emissions to avoid double counting. (We further discuss related approaches and provide an overview in the SM and table S1.)

It is conceptually important to underscore that our approach to expressing carbon damages as a fraction of operating profits provides neither the percentage by which profits would decline if carbon emissions were taxed at the SCC (*6*) nor a measure of the implicit subsidy to firms' profits due to insufficient carbon regulation globally (*7*). The impact of a carbon tax (set at the SCC) on profits would differ from corporate carbon damages because the tax's incidence would be divided between firms, customers, and workers. The exact split would depend on several factors, including the elasticities of supply and demand for firms' products, which vary greatly across and within industry and country (*3*). Also, carbon damages are not equivalent to and exceed the implicit carbon subsidy of corporate activities because firms pay for emissions in some regimes (e.g., the EU Emissions Trading System) and, in addition, face nonprice emissions regulation in many parts of the world. Although the explicit or implicit carbon prices to firms from these regulations are generally much lower than the SCC, one would need to measure the per-ton cost of carbon regulation (price and nonprice) that firms' emissions face globally to compute subsidies. The data necessary to do such calculations are not available.

Thus, our analysis is a "first cut" of the corporate carbon damages that is based on one of the largest emissions datasets that is presently available. It serves as a preview for what would be revealed with global mandatory reporting. We underscore, however, that the resulting estimates must be interpreted cautiously. Of the firms in our sample for which GHG emissions data were available, only 31% of them directly reported their emissions (fig. S1). This reflects that in almost all parts of the world, reporting is, at present, still voluntary, lacks independent verification, and/or faces no penalties for underreporting; together, these features raise important questions about the reliability of emissions data that we cannot answer. The emissions for the remaining 69% of the sample firms are estimated by Trucost using a model that relies on mostly voluntary emissions reports from a wide array of data sources to determine sector-specific emission intensities, the company's business sectors, and its revenue share by sectors (see SM). The results are qualitatively unchanged when using a subset of emissions data that is reported as third-party verified or when using an alternative data provider with less coverage (table S14).

Corporate carbon damages by country

The table presents estimates of average corporate carbon damages by country as well as country rankings based on the (unadjusted) damages and industry-adjusted corporate carbon damages.

COUNTRY	*N*	UNADJUSTED MEAN (%)	UNADJUSTED RANKING	INDUSTRY-ADJUSTED RANKING
Australia	293	35.6	13	14
Brazil	148	41.0	9	16
Canada	297	38.0	10	15
China	1883	56.3	5	5
European Union	740	37.9	11	12
France	213	29.5	15	13
Germany	201	42.5	8	6
India	524	78.8	3	7
Indonesia	134	89.6	2	2
Italy	114	36.7	12	10
Japan	2149	30.7	14	4
Mexico	80	67.0	4	9
Russia	55	129.6	1	3
South Africa	122	50.7	6	8
South Korea	726	45.8	7	1
United Kingdom	385	21.7	17	17
United States	2091	25.7	16	11

Carbon damages are calculated with a social cost of carbon of $190 per ton of CO_2-equivalent emissions and the global public firm sample and are expressed as a percentage of operating income. The country sample is restricted to those listed by the Major Economies Forum on Energy and Climate. This list includes the European Union (EU) as well as three of its member states (France, Germany, and Italy). We report these countries separately and combine the remainder of the EU countries. The industry adjustment is performed by regressing firm-level damages on a complete set of binary industry indicators (see the second figure for industries). We use firm-level damages scaled by operating income and apply the natural logarithm given the skewed distribution. Rank 1 (17) indicates the highest (lowest) average carbon damages. We keep only observations for which operating income is positive. We truncate observations of the scaled carbon damages that are below the first or above the 99th percentile.

RESULTS

In presenting the results, we focus on the global sample and additionally highlight the subsample of US companies. Aside from being the largest economy and home to a large set of publicly traded firms, the United States plays a special role because of its current regulatory debate over mandatory climate reporting. We also focus on the damage estimates when the SCC is $190 per tCO_2e. (Estimates using other SCC values are provided in the SM.)

Three main findings emerge from our analysis. First, average corporate carbon damages are large, but they vary greatly across firms. For the global sample, they equal roughly 44% of firms' operating profits and 3.1% of their revenues. For US firms, average damages are 18.5% of profits and 2% of revenues (see the first figure and tables S3 and S4). When we calculate corporate carbon damages as a weighted average to account for firm size differences, using operating profits as weights, the global and the US averages are 34.2 and 15.6% of profits, respectively (fig. S2).

Second, there is substantial variation across firms, as well as across and within industries. The means far exceed the medians because there are firms with outsized carbon damages (see the first figure). But it is important to recognize that there are firms on both ends of the spectrum. In the global sample, the 10th percentile of carbon damages is equal to only 0.1% of corporate profits, whereas it is 85.8% for the 90th percentile (table S3).

Next, the largest carbon damages occur in the energy-intensive industries (i.e., utilities, materials, energy, transportation, and food, beverage, and tobacco), for which the industry average of the damages is well above the mean of the global sample (see the second figure and fig. S3). It is noteworthy that the top-four industries account for 89% of the total global corporate carbon damages. Importantly, however, there is substantial heterogeneity within industry. For example, globally (in the United States) corporate carbon damages as a share of profits in the energy sector are 382.9% (233.7%) for the 90th-percentile firm and just 4.5% (4.5%) for the 10th-percentile firm in this sector (tables S5 and S6).

This heterogeneity within the same industry suggests that peer benchmarking has the potential to induce meaningful reductions in corporate carbon damages (table S13). To illustrate this potential, we compute the amount by which firms' total emissions would decline, if all firms with carbon damages above their industry's median were to reduce their carbon damages to their respective industry median. For our sample, total emissions would decrease by more than 70%, with either the operating profit or the revenue normalization. Although this is just one example, it makes clear that benchmarking firms against their peers in terms of carbon damages has high potential for reducing corporate emissions (more details in table S12).

Third, the variation in carbon damages across countries is substantial (fig. S4 and tables S7 and S8). Countries with high unadjusted average damages are Russia (129.6%), Indonesia (89.6%), and India (78.8%). The average unadjusted dam-

ages for France, the United States, and the United Kingdom are 29.5, 25.7, and 21.7%, respectively (see column 2a of the table). The unadjusted rankings accord roughly with conventional wisdom about differences in climate regulatory stringency (see column 2b of the table). They also reflect carbon production, rather than consumption; so, for example, the climate damages from high–carbon intensity products (e.g., steel bars) that are imported into the United States are assigned to the exporter, rather than to the United States.

The country rankings are influenced by differences in industrial composition (see column 2b of the table), so the country rankings after adjustment for these differences are reported (see column 3 of the table). This normalization changes the country ranking. For instance, Brazil has relatively many firms in carbon-intensive sectors, such as transportation and utilities, and thus moves down from 9th to 16th place when we account for its industrial composition. Conversely, South Korea has relatively many firms in less–carbon intensive industries, such as semiconductors, and hence rises from 7th place to being the country with the highest average industry-adjusted corporate carbon damages.

CONCLUSION

The core finding from our analysis is that corporate carbon damages are, on average, large but highly skewed, with median damages being much smaller. Moreover, these damages are heterogeneous across industries as well as within industries and countries. It is important to bear in mind that these findings are largely derived from voluntary reporting with no penalties for misreporting or even from estimated emissions. This is not a small caveat and underscores the need for mandatory and verified emissions reporting.

Mandatory disclosure can aid in decreasing GHG emissions in at least three important ways. First, it is not possible to have meaningful policies that aim to restrict GHG emissions without reliable measurement and credible data. This is true for both market-based policies (e.g., taxes on GHG emissions and cap-and-trade markets) and for command-and-control policies, which also require credible data to determine whether the policy is achieving its intended goals.

Second, mandatory disclosure would help financial markets to discipline GHG emissions by pricing existing and expected future environmental policies. Such disclosure and the subsequent pricing would also give firms incentives to think strategically about their GHG emissions. Supporting this view, a considerable body of research suggests that financial disclosure mandates have improved market pricing of risks, capital allocation, and firms' financial operations (*1*, *8*, *9*), and indeed they are the bedrock of capital markets.

Third, recent studies show that disclosure mandates can incentivize firms to reign in environmental externalities, such as GHG emissions, even in the absence of environmental policy [e.g., (*10*, *11*)]. Targeted transparency has been used successfully in other policy areas (*12*). Thus, there is an empirical basis for the view that mandatory disclosure could pressure firms to reduce their GHG emissions. At the same time, we note that this "channel" relies on the nonfinancial preferences of key stakeholders (e.g., employees, customers, and perhaps even shareholders), which expands firms' social responsibility beyond profit maximization.

Carbon emissions, such as from this coal-fired power plant in Pennsylvania, USA, are released into the atmosphere, leading to global climate damages.

A potentially dispositive weakness in firms' claims to achieve "net zero" and other promises about future GHG emissions reductions is the availability of reliable data regarding whether firms are living up to their promises or engaging in "greenwashing" that does not produce real emissions reductions. Mandatory disclosure would provide a way to hold firms accountable for their promises by providing annual assessments of their own and their competitors' progress. Such benchmarking against previous years' emissions or peer firms' emissions could unlock continued emissions reductions (*10*). However, to be successful, emissions disclosures have to be credible, and the regime should ideally cover all but the smallest private and public firms.

We believe that Supreme Court Justice Louis Brandeis's famous 1913 prescription that "sunlight is the best disinfectant" for "social and industrial diseases" still has merit more than a century later in confronting the climate challenge, a problem that society was unaware of at that time. Revealing corporate carbon damages would start a public dialogue about the contribution of corporate activities to the climate problem, which in turn could spur policies and unleash market forces. Put plainly, it is difficult to imagine a successful approach to the climate challenge that does not have widespread mandatory disclosure as its foundation. ■

REFERENCES AND NOTES

1. C. Leuz, P. Wysocki, *J. Account. Res.* **54**, 525 (2016).
2. T. Tietenberg, *Environ. Resour. Econ.* **11**, 587 (1998).
3. M. J. Kotchen, *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2011969118 (2021).
4. S&P Global, "Trucost environmental register: Methodology FAQs" (S&P Global, 2019); www.jpx.co.jp/corporate/sustainability/esgknowledgehub/esg-rating/nlsgeu0000053wxn-att/Trucost_Environmental_Register_Methodology_FAQs.pdf.
5. T. Carleton, M. Greenstone, *Rev. Environ. Econ. Policy* **16**, 196 (2022).
6. R. Eccles, X. Gao, S. Rajgopal, "How a carbon tax would hit the earnings of US companies," *Responsible Investor*, 1 March 2022.
7. D. Coady, I. Parry, L. Sears, B. Shang, *World Dev.* **91**, 11 (2017).
8. S. Roychowdhury, N. Shroff, R. S. Verdi, *J. Account. Econ.* **68**, 101246 (2019).
9. M. Greenstone, P. Oyer, A. Vissing-Jorgensen, *Q. J. Econ.* **121**, 399 (2006).
10. S. Tomar, *J. Account. Res.* **61**, 451 (2023).
11. P. Bonetti, C. Leuz, G. Michelon, "Internalizing externalities: Disclosure regulation for hydraulic fracturing, drilling activity and water quality," Working paper 30842, National Bureau of Economic Research, January 2023; https://www.nber.org/system/files/working_papers/w30842/w30842.pdf.
12. D. Weil, M. Graham, A. Fung, *Science* **340**, 1410 (2013).
13. M. Greenstone, C. Leuz, P. Breuer, GLB-2023: Initial release. Zenodo (2023); https://doi.org/10.5281/zenodo.8187694.

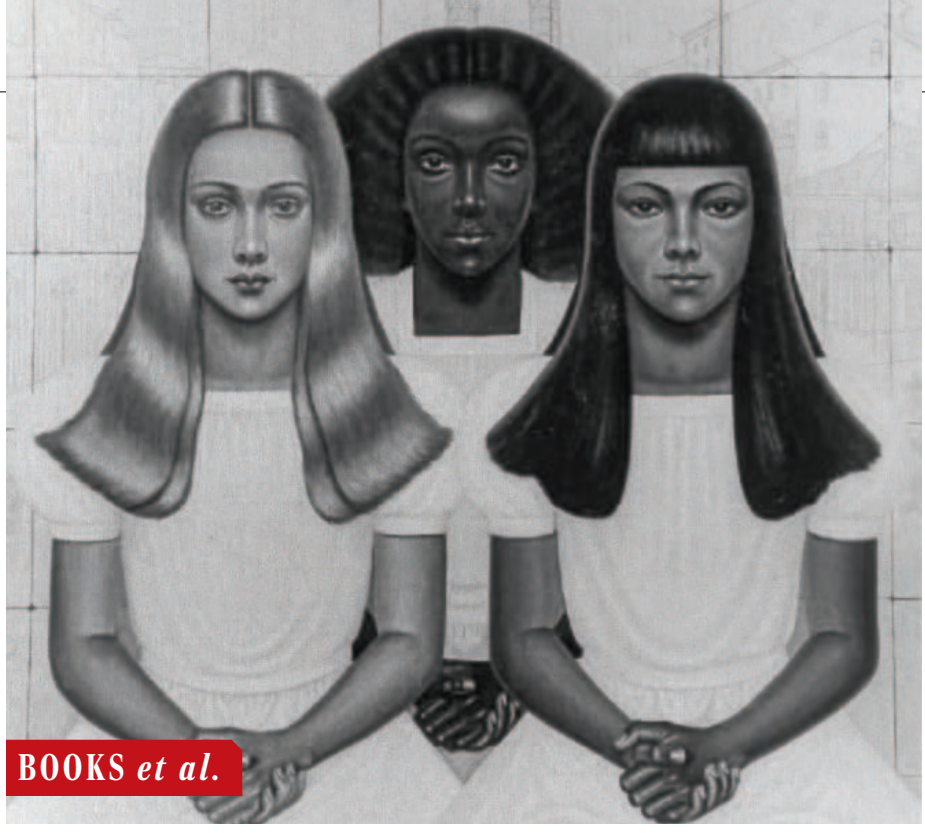
ACKNOWLEDGMENTS

The authors thank K. Ito, C. Knittel, C. Wolfram, and T. Sellhorn for comments. M.G. acknowledges research funding from the University of Chicago Energy Policy Institute (EPIC). C.L. acknowledges research funding from the Rustandy Center and the Booth School of Business of the University of Chicago. P.B. acknowledges research funding from the German Research Foundation (DFG) under the TRR 266 program ("Accounting for transparency") and financial support from the Julius Paul Stiegler Memorial Foundation, Landesgraduiertenförderung Baden-Württemberg, and the German Academic Exchange Service (DAAD). M.G. is a cofounder of Climate Vault, a registered nonprofit 501(c)(3) organization, and chairman of the board for the for-profit company Climate Vault Solutions. M.G. owns a diversified portfolio that includes companies in the paper's underlying dataset. C.L. and P.B. have no competing interests. Raw data were obtained from Trucost and S&P Capital IQ Pro by license agreements for noncommercial use. More details about the analyses are provided in the supplementary materials and in (*13*).

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.add6815

10.1126/science.add6815



This illustration by Dimitri Ismailovitch appeared in Ramos's *Introduction to Brazilian Anthropology*.

ILLUSTRATION: DIMITRI ISMAILOVITCH

HISTORY OF SCIENCE

A history of anti-racism in science

Twentieth-century scientists were imperfect diversity advocates in the Global South

By **Jörg Matthias Determann**

On 10 June 2020, researchers in the United States and elsewhere joined the Strike for Black Lives, a mass walkout that sought to draw attention to systemic racism and racial inequality. This high-profile protest might have given the impression that anti-racism in science is a relatively recent concern, propelled by the Black Lives Matter movement that emerged in the 2010s. However, just as feminism in academia predates the #MeToo movement, anti-racism in science has a much longer history.

In 1950, at the beginning of the Cold War, the United Nations Educational, Scientific and Cultural Organization (UNESCO) issued the first in a series of statements on race, condemning racism and emphasizing the unity of the human species and the equality of all its members. Previous historians have understood UNESCO's campaign against race prejudice as mainly a crusade undertaken in the North Atlantic world, with British and US researchers being seen as the moral leaders who responded to the horrors of scientific racism in Nazi-occupied Europe. In his new book, *The Remnants of Race Science*, historian Sebastián Gil-Riaño refocuses this story on the Global South and especially on Latin America.

The protagonists of Gil-Riaño's book are scientists who spent much of their careers examining non-European communities, including Indigenous people in Peru and descendants of African slaves in Brazil. These researchers sought to replace an emphasis on biological differences with a concern for cultural diversity.

The central character of Gil-Riaño's book is Alfred Métraux, a Swiss-born anthropologist who studied the native ethnic groups of Argentina and served as director of UNESCO's race campaign in the 1950s. Rejecting ideas about racial order based on biological differences that were common at the time, Métraux believed that "given similar degrees of cultural opportunity to realize their potentialities, the average achievement of the members of each ethnic group is about the same."

Other activist scholars pushed back against racist attitudes too. The Brazilian psychologist Arthur Ramos, for example—another prominent figure in Gil-Riaño's book—saw criminality among Black and mixed-race people in his home country as being due to social factors rather than representing inherited traits. Ramos would become the head of UNESCO's Social Sciences Department at the start of the organization's campaign to publicize scientific understandings of race and to condemn racism.

Drawing on primary sources recorded in Spanish, Portuguese, and French, Gil-Riaño introduces the reader to many scientific figures and projects that are relatively unknown in Anglophone spheres. Ramos, we learn, for example, was preceded by Raimundo Nina Rodrigues, a doctor of legal medicine who was a foundational figure in the field of Afro-Brazilian studies. Rodrigues's scholarship inspired Ramos and others to advocate for the United Nations to embark on projects to improve conditions for nonwhite communities that were perceived to be "primitive."

Even if mid-20th-century scientists in the Global South made advances in the struggle against race prejudice, their attitudes and beliefs were still shaped by their times. Although they laid to rest some of the most extreme eugenic ideas of biology as destiny and championed ideals of equality and "unity in diversity," many were unable to escape other biases and power asymmetries that still structured international relations after World War II. While they did not consider Indigenous and non-European people to be members of inferior races, for example, many felt that such individuals were held back by their cultures. A hierarchy of a modern "West" above a backward "Rest" thus continued to hold sway for much of the 20th century.

The Remnants of Race Science
Sebastián Gil-Riaño
Columbia University Press, 2023. 392 pp.

Gil-Riaño's rich and detailed account of Southern scientists' efforts to push back against pervasive racism thus offers a cautionary tale rather than a blueprint for today's anti-racist scholars. Nonetheless, his book offers useful historical context to current debates about how to successfully build solidarity in science and society. ■

The reviewer is at the Liberal Arts and Sciences Program, Virginia Commonwealth University School of the Arts in Qatar, Doha, Qatar. Email: jmdetermann@vcu.edu

10.1126/science.adi3191

SCIENCE AND SOCIETY

Embrace the improbability of existence

Learning to appreciate the rarity of life could improve our climate future, argues a physicist

By **Matthew Stanley**

In *The Dawn of a Mindful Universe*, physicist Marcelo Gleiser attempts to save the world by way of a pessimistic but spiritual application of the Drake equation—the mathematical framework used to estimate the number of intelligent civilizations likely to exist in the Milky Way Galaxy. Gleiser has written a number of stimulating books on the existential implications of modern science, and here he continues that trajectory in what he calls a manifesto of "urgency and hope." His goal is to change humanity's collective mindset to divert us from the "delusional and suicidal" climate path on which we find ourselves by reframing the stories we tell ourselves about the Universe and about humankind.

The Dawn of a Mindful Universe is a wide-ranging, fast-paced journey from the pre-Socratics to *Star Trek*. It is both a story of the evolution of the Universe and life and a story of our understanding of them—a very wide scope for a relatively short book, which means that not every topic gets equal coverage or analysis. At times, this leads to some unevenness, but it is a worthwhile price to pay for gems such as Gleiser's insightful discussion of the multiverse as a secular "God of the gaps."

Gleiser's overall goal is to chart how we came to value ourselves over nature and, in parallel, came to think that the Universe should be full of life-forms like us. He argues instead for a "post-Copernican" mindset, one that embraces the preciousness of our planet over its similarities to other worlds. Here, he builds on his earlier writings about the limitations of science to cast doubt on the uniformity principles used to assume that life is common in the Universe, warning readers not to extrapolate what we know about life on Earth to the rest of the cosmos. His arguments echo those made by paleontologist Stephen Jay Gould, who stressed the contingency of life's emergence and development, arguing that if a single event had been different, we might not be here.

The reviewer is at the Gallatin School of Individualized Study, New York University, New York, NY 10003, USA.
Email: ms5100@nyu.edu

These discussions build up to Gleiser's conservative reading of the Drake equation. In recent years, many scientists have been leveraging discoveries in exoplanets and astrobiology to give generous estimates of the frequency of intelligent life in our Galaxy. Gleiser disagrees strongly with this approach.

Going against the tide of opinion, Gleiser argues that the failure so far to find life elsewhere means that we are, essentially, the only intelligent beings in the Milky Way—a conclusion he uses to call for a resacralization of nature. We should embrace a biocentric view that life must be protected as something unique and endangered, he maintains.

There is no supernaturalism here. Gleiser takes a nondenominational position that allows and encourages transcendent, sublime experiences of life. He rejects physicist Steven Weinberg's classic conclusion about the apparent pointlessness of the Universe, instead contending that the more we understand the cosmos, the more we understand our own rarity and therefore preciousness.

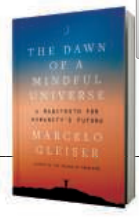
Gleiser suggests combining this cosmological interpretation with a kind of spirituality inspired by Indigenous traditions in which nature is revered. Like historian Lynn White Jr., he contends that the root of our ecological crisis is religious in nature and the solution must therefore be rooted in science-driven spirituality that establishes interconnectedness within the biosphere.

The book's closing chapter presents a six-point plan to change the course of civilization by embracing a new conception of humanity as a single tribe with shared values. As is often the case with manifestos, though, many of Gleiser's points are simply stated rather than argued for. The rarity and interconnectedness of life may well be supported by scientific evidence, but the moral responsibility Gleiser hopes will follow is not. As philosophers sometimes point out, it is difficult to go from "is" to "ought." Other resources are generally needed to establish value systems. Such a role might be played by the Indigenous systems that Gleiser attempts to bring into the conversation, although they are not discussed in much detail in the book.

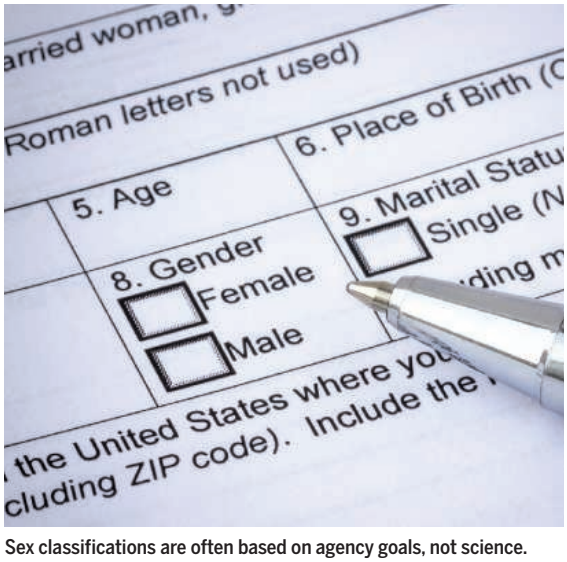
As with the long-standing tradition of natural theology, it seems unlikely that *The Dawn of a Mindful Universe* is going to convince anyone who does not already agree with Gleiser's position that the biosphere needs protecting. However, it will no doubt provide ammunition and many interesting ideas to those who already believe. ■

10.1126/science.adj0232

The Dawn of a Mindful Universe: A Manifesto for Humanity's Future
Marcelo Gleiser
HarperOne, 2023. 256 pp.



SCIENCE, SEX, AND GENDER



Sex classifications are often based on agency goals, not science.

PODCAST

Sex Is as Sex Does: Governing Transgender Identity
Paisley Currah
New York University Press, 2022. 256 pp.

When and how governments assign sex classifications to citizens is less straightforward than many might think and is more often related to social and political goals than to scientific understandings of sex. This week on the *Science* podcast, Paisley Currah discusses the myriad ways agencies of the state navigate sex classification and how efforts to label people in this way can be problematic for all, not just for individuals who do not easily fit into existing categories.
bit.ly/3PX6dNU

10.1126/science.adj8037

PHOTO: UDAYAN BORTHAKUR

Edited by **Jennifer Sills**

Rerouting railway would conserve India's gibbons

The western hoolock gibbon (*Hoolock hoolock*), an obligate canopy dweller that rarely ventures to the ground, inhabits dense evergreen and semi-evergreen forests in south and southeast Asia. The globally endangered primate species is threatened by habitat destruction and fragmentation as well as hunting (*1*). Even in the Hollongapar Gibbon Sanctuary in Assam, India, a 21-km^2 protected area with about 125 gibbons from an identified "northern" subpopulation (*1*, *2*), the lowland rainforest habitat is fragmented by a railway track. To ensure that the gibbon sanctuary provides effective conservation, the track should be rerouted.

India harbors at most 10,000 western hoolock gibbons (*3*). Their small, insular populations are restricted to the southern and eastern banks of Brahmaputra and Dibang rivers, respectively (*1*). Although India has given the western hoolock gibbon the highest legislative protection (*4*), no national or state-level conservation action plan for the species exists.

A railway track has divided the habitat that is now the Hollongapar Gibbon Sanctuary since 1887 (*5*). Between 1880 and 1920, human settlements and tea plantations, established by the then British colonial government, severed the area from surrounding forests (*1*). The state government of Assam made the sanctuary a protected area in 1997 (*5*). Additional land-use changes to and forest loss and degradation in the surrounding areas in recent decades have left an isolated "forest island" (*1*, *6*). Because the western hoolock gibbon is sensitive to canopy gaps and habitat fragmentation, the movements and dispersal of the species have become severely limited in the sanctuary and surrounding region (*5*), potentially compromising this subpopulation's long-term genetic and demographic viability (*2*).

Recently, the Assam Forest Department and Wildlife Institute of India jointly finalized a detailed design of artificial canopy bridges to help reconnect gibbon groups on either side of the existing railway track (*5*). Initiatives are also underway to restore connectivity between the sanctuary and surrounding forests (*6*). Elected representatives have indicated support for gibbon conservation (*7*, *8*). However, impending track electrification and expansion plans (*9*, *10*) pose insurmountable challenges to such conservation efforts.

The Indian government promotes sustainable development and emphasizes coexistence with wildlife and habitats in its growth philosophy (*11*, *12*). Given the western hoolock gibbon's conservation status, Hollongapar Gibbon Sanctuary's small size, and the availability of surrounding nonforest land, the existing railway track should be rerouted to areas outside the sanctuary. Moving the track would be a critical step toward gibbon conservation and would also unambiguously demonstrate the government's commitment to balancing India's rapidly growing economy with its fragile ecology.

Rohit Ravindra Samita Jha[1], Gopi Govindan Veeraswami[1]*, Dilip Chetry[2,3], Nandha Kumar[4]
[1]Wildlife Institute of India, Dehradun, Uttarakhand, India. [2]Aaranyak, Guwahati, Assam, India.

Development in India's Hollongapar Gibbon Sanctuary threatens the western hoolock gibbon population.

[3]International Union for Conservation of Nature Species Survival Commission Primate Specialist Group, Guwahati, Assam, India. [4]Department of Environment and Forests, Government of Assam, Jorhat, Assam, India.
*Corresponding author. Email: gopigv@wii.gov.in

REFERENCES AND NOTES

1. D. Chetry, R. Chetry, A. K. Das, R. Bhuyan, R. C. Kyes, *Prim. Conserv.* **36**, 233 (2022).
2. M. Trivedi, K. D. Sree, J. Biswas, G. Umapathy, *Int. J. Primatol.* **44**, 1 (2023).
3. Global Gibbon Network, "Gibbons by Country (NE India)" (2022); http://www.globalgibbonnetwork.org/neindia.
4. Wild Life (Protection) Act, Union of India, Act 53 of 1972 (1972); https://www.indiacode.nic.in/bitstream/123456789/1726/1/a1972-53.pdf.
5. R. R. S. Jha, S. Zangmo, G. V. Gopi, "Artificial Canopy Bridge Design to facilitate Western Hoolock Gibbon (*Hoolock hoolock*) crossing over Mariani-Dibrugarh broad gauge single track Railway Line in Hollongapar Gibbon Sanctuary, Assam" (TR No/2023/07, Wildlife Institute of India, Dehradun, 2023); https://wii.gov.in/images/images/documents/publications/rr_2023_canopy-bridge_design-guidelines.pdf.
6. Saving Nature, "India's gibbons need more forest" (2019); https://savingnature.com/india-hollongapar-gibbons-sanctuary.
7. Office of the Chief Minister of Assam, 12th State Board for Wild Life (SBWL) meeting, Assam, "Minutes of meeting" (2022); https://forestsclearance.nic.in/writereaddata/Sec_Letter/0_0_07_Sep_2022_122429013_minuts_Minutesofthe12thMeeting.pdf.
8. Ministry of Environment, Forest, and Climate Change, Government of India, "Reply to starred question no. 267 asked by Member of Parliament (Lok Sabha) Mr. Topon Kumar Gogoi (17th Lok Sabha, Session no. 11), Jorhat Parliamentary Constituency" (2023); https://sansad.in/getFile/loksabhaquestions/annex/1711/AS267.pdf?source=pqals.
9. Northeast Frontier Railway, Indian Railways, "NFR undertaken various developmental works in Assam: Aiming better rail connectivity and passenger comfort" (2022); https://nfr.indianrailways.gov.in/view_detail.jsp?lang=0&id=0,4,268&dcd=1898.
10. Northeast Frontier Railway, Indian Railways, "N. F. Railway General Manager conducts review meeting with MPs: Various aspects of Railway Developmental plans in Tinsukia division highlighted" (2022); https://nfr.indianrailways.gov.in/view_detail.jsp?lang=0&dcd=2006&id=0,4,268.
11. Press Information Bureau, Ministry of Information and Broadcasting, Government of India, "Text of PM's address at the inauguration of 13th COP to the Convention on Migratory Species in Gandhinagar Gujarat" (2020); https://pib.gov.in/PressReleaseIframePage.aspx?PRID=1603380.
12. Press Information Bureau, Ministry of Information and Broadcasting, Government of India, "Shri Bhupender Yadav says that under India's G20 Presidency, the concept of 'Green Growth' being a priority area validates how sustainable development has been mainstreamed in Indian policy making process through a visionary approach" (2023); https://pib.gov.in/PressReleasePage.aspx?PRID=1901475.

10.1126/science.adj5002

Norway's mining plans threaten Arctic life

The demand for metals is surging in response to the rapid development of green energy infrastructure such as electric vehicle

0825_Letters_17691295.indd 843 8/21/23 3:58 PM

batteries and magnets (*1*). As a result, interest in deep-sea mining for critical minerals is increasing (*2*), potentially putting species—many of which have only recently been identified (*3*)—at risk of pollution (*4*) and marine habitat destruction (*2*). Data on benthic and pelagic biodiversity and deep-sea ecosystem dynamics is inadequate, hindering essential environmental impact assessments (*5*, *6*) that could quantify the impacts of deep-sea mining. Despite these uncertainties, Norway has proposed deep-sea mining on its continental shelf.

Norway's mining plans cover an area of about 280,000 km^2, including areas of the Norwegian and Greenland Seas (*7*). The total area is estimated to contain about 38 million tons of copper, which is more than the total global annual extraction (*8*). The proposed mining activities will threaten deep-sea keystone species and wildlife (*2*, *5*, *6*).

Elsewhere in Europe, deep-sea mining has been put on hold. France has banned it, and Germany has called for a "precautionary" pause while awaiting better data on its potential impacts on biodiversity and ecology (*9*). The Norwegian parliament is scheduled to vote on deep-sea mining later this year (*7*). If approved, Norway would become one of the first nations to initiate such operations.

Norway already has a large direct (*10*) and indirect carbon footprint per capita because of its extensive oil exploitation, threatening the vulnerable Arctic region (*11*). The Norwegian government should carefully consider the entire spectrum of environmental risks associated with deep-sea mining before approving the proposed projects. Deep-sea mining could lead to losses that surpass critical planetary tipping points, causing irreversible damage to ecosystem services, including but not limited to fisheries. Instead of mining vulnerable seabed ecosystems, Norway should increase the availability of metals by initiating better global recycling (*5*).

Christian Sonne[1,2]*, Tomasz M. Ciesielski[3,4], Bjørn M. Jenssen[1,3,4], Su Shiung Lam[5], Huan Zhong[6], Rune Dietz[1]

[1]Department of Ecoscience, Aarhus University, DK-4000 Roskilde, Denmark. [2]University of Petroleum and Energy Studies, Uttarakhand, India. [3]Department of Biology, Norwegian University of Science and Technology, Trondheim, Norway. [4]Department of Arctic Technology, The University Centre in Svalbard (UNIS), NO-9171 Longyearbyen, Norway. [5]Universiti Malaysia Terengganu, Terengganu, Malaysia. [6]School of Environment, Nanjing University, Nanjing, China.
*Corresponding author. Email: cs@ecos.au.dk

REFERENCES AND NOTES

1. G. Calvo, A. Valero, *Environ. Dev.* **41**, 100640 (2022).
2. D. J. Amon *et al.*, *Mar. Policy* **138**, 105006 (2022).
3. M. Rabone *et al.*, *Curr. Biol.* **33**, 2383 (2023).
4. N. C. Mestre, R. Calado, A. M. V. M. Soares, *Environ. Pollut.* **185**, 369 (2014).
5. K. Bisson, H. McMonagle, I. Iglesias, S. Halfter, N. Gallo, *Commun. Earth Environ.* **4**, 152 (2023).
6. S. Katona *et al.*, *Biodivers. Conserv.* **32**, 1125 (2023).
7. K. Bryan, R. Milne, "Norway seeks to open vast ocean area to deep-sea mining," *Financial Times*, 9 June 2023.
8. Norwegian Petroleum Directorate, "Resource assessment of seabed minerals" (2023); https://www.npd.no/globalassets/1-npd/fakta/havbunnsmineraler/publikasjoner/2023/ressursvurdering-havbunnsmineraler-20230127.pdf [in Norwegian].
9. Environmental Justice Foundation, "Germany, France and Spain create genuine possibility for action on deep sea mining" (2022).
10. H. Ritchie, M. Roser, "Norway: CO_2 country profile," Our World in Data (2020).
11. N. Adomaitis, G. Fouche, "Norway plans to offer record number of Arctic oil, gas exploration blocks," *Reuters*, 24 January 2023.

10.1126/science.adj4244

Norway's mining in Arctic waters would put the region's species and ecosystems at increased risk.

Threat to Tanzanian Park requires compromise

Katavi National Park in western Tanzania is famed for its pristine remoteness and large mammal dry season aggregations (*1*, *2*). Two unpaved roads of equivalent length run through the park, linking the regional capitals of Mpanda and Sumbawanga. One road cuts across the center of the park, crossing the Katuma River, a dry season hotspot for very large mammals, especially hippopotamus (*3*). The other runs farther west through miombo woodland with much lower animal densities. After a hiatus due to lack of funding, construction to pave the road that bisects the central area frequented by animals in the dry season has recently started, putting animal populations at risk. Yet Tanzania could still craft a win-win solution to this environment-development conflict if it reassesses its paving decision.

The road upgrade will lead to the loss of at minimum 188,000 trees along the 46-km section through the park (*4*), an increase in road collisions between cars and large mammals (*5*), and numerous roadkills of amphibians, birds, and reptiles (*6*). A tarmac road will further undermine public perception of the area as a protected wilderness, thereby reducing lucrative tourism revenue generated principally by the national park's remoteness [e.g., (*7*)].

Tanzania National Parks (TANAPA), which is relatively powerless against the government authority responsible for road building and maintenance in Tanzania (TANROADS), plans to reduce animal mortality by authorizing speed bumps on the road and placing gates at the park boundaries. These measures may offer limited protection but will likely be unable to counter the expected increase in traffic volume due to drivers choosing the upgraded paved road over the unpaved murram road, drivers racing between speedbumps, noise, and littering (*4*).

In another Tanzanian national park, Mikumi, a paved road led to high rates of animal–vehicle collisions involving human casualties, vehicular damage, and wildlife mortalities (*5*). In response, paved roads through national parks were deemed unwise (*8*) and plans were revoked in Serengeti (*9*). Although the practice has restarted, a compromise solution is still possible (*10*).

Extensive tree felling to widen the Katavi road is already in full swing, but the Tanzanian Government could stop paving 5 miles (8 km) south and 5 miles north of the Katuma River, giving elephants, buffalo, hippopotamus, and giraffes the freedom to graze and browse relatively undisturbed in proximity to dry season aggregation areas. Urban connectivity and Tanzania's Natural Capital could both be sustained if the Government bravely modifies its decision and acts with agility and speed.

Tim Caro[1]*, Andrew Dobson[2], Christian Kiffner[3], Matthias Waltert[4]

[1]School of Biological Sciences, University of Bristol, Bristol BS8 1TQ, UK. [2]Ecology and Evolutionary Biology, Princeton University, Princeton, NJ 08544, USA. [3]Leibniz Centre for Agricultural Landscape Research, 15374 Müncheberg, Germany. [4]Department of Conservation Biology, University of Gottingen, 37073 Gottingen, Germany.
*Corresponding author.
Email: tim.caro@bristol.ac.uk

REFERENCES AND NOTES

1. C. Stuart, M. Stuart, *National Parks and Game Reserves of East Africa (Struik Nature Field Guides)* (Penguin Random House South Africa, 2019).
2. T. Caro, *J. Appl. Ecol.* **36**, 205 (1999).
3. M. Waltert *et al.*, *Afr. J. Ecol.* **47**, 737 (2009).
4. T. Caro, *Trop. Conserv. Sci.* **8**, 1009 (2015).
5. C. Drews, *Afr. J. Ecol.* **33**, 89 (1995).
6. J. Kioko *et al.*, *Afr. Zool.* **50**, 17 (2015).
7. Chada Katavi (2023); https://www.nomad-tanzania.com/camps/western-tanzania/katavi-national-park/chada.
8. W. F. Laurance *et al.*, *Curr. Biol.* **25**, R259 (2015).
9. A. P. Dobson *et al.*, *Nature* **467**, 272 (2010).
10. T. Caro *et al.*, *Curr. Biol.* **24**, R722 (2014).

COMPETING INTERESTS

A.D. is vice chair of Frankfurt Zoological Society–United States, a voluntary position.

10.1126/science.adj4393

RESEARCH

IN *SCIENCE* JOURNALS

Edited by **Michael Funk**

PHOTO: NIGEL CATTLIN/MINDEN PICTURES

Wild ancestors of potatoes are able to sense and resist oomycete pathogens that cause late blight, pictured here on a domesticated potato leaf.

PLANT SCIENCE

Potato pathogen resistance

The oomycete *Phytophthora infestans* devastates potato crops, most famously during the Great Irish Famine of the mid-1800s. Torres Ascurra *et al.* examined wild potato variants from across the Americas and identified a pattern recognition receptor called PERU, which recognizes a *P. infestans* peptide. When PERU binds a protein fragment from *P. infestans*, the potato plant can mount an immune response. The authors established that different alleles of PERU from wild Andean potato relatives have different sensitivities to the *P. infestans* peptide. Their work provides mechanistic insight into *P. infestans* immunity, thus paving the way for improved crop resilience to a disease that has been challenging to control. — MRS *Science*, adg5261, this issue p. 891

CATALYSIS

Cooling down methane pyrolysis

Very high temperatures are needed to convert methane to hydrogen and solid carbon. A known liquid metal nickel–bismuth catalyst for this pyrolysis reaction resists deactivation but requires reaction temperatures above 1000°C. Chen *et al*. added molybdenum to improve nickel reactivity, which could enable lower operating temperatures. Molybdenum appears to decrease the negative charge on nickel atoms in the molten state and to increase atom mobility and interaction with methane, which enables efficient long-term pyrolysis at 800°C. —PDS

Science, adh8872, this issue p. 857

ORGANIC CHEMISTRY

Displacing olefins with nitrogen

Reactions that form carbon–nitrogen bonds most often target carbon centers that are either single bonded to a halogen or double bonded to oxygen or another carbon. He *et al*. present an alternative sequence that targets a carbon–carbon single bond adjacent to an olefin. Treatment of the allylic carbon compound with ozone followed by copper catalysis formally displaces the pendant olefin with an amine. The reaction can introduce nitrogen into a wide variety of complex terpenes, among other compounds. —JSY

Science, adi4758, this issue p. 877

CHEMICAL TECHNOLOGY

Indirect oxidative propylene synthesis

The conversion of propane to propylene by oxidation is less energy intensive than direct dehydrogenation but must be controlled to avoid overoxidation, which limits selectivity and conversion. Wang *et al*. show that a chemical looping process in which the feed gas switched between propane and oxygen achieved 81.3% propylene selectivity at 42.7% propane conversion at 550°C. Well-dispersed vanadium dioxide on alumina performed the dehydrogenation reaction, and adjacent iron vanadate particles acted as the oxygen carrier. Propylene selectivity increased with nanoscale proximity of the catalyst sites so that the hydrogen generated at the vanadium dioxide could migrate readily to the adjacent iron vanadate side for combustion. —PDS

Science, adi3416, this issue p. 886

ALSO IN *SCIENCE* JOURNALS

Edited by **Michael Funk**

CANCER GENETICS

Not just random chromosome changes

Aneuploidies, which are changes in the numbers of whole chromosomes or chromosome arms, are common in cancer, but their contributions to cancer cell survival have been difficult to pinpoint. Girish *et al*. developed a chromosome-engineering tool to orchestrate the targeted loss of aneuploid chromosome arms and thereby compare isogenic cancer cell lines with and without selected trisomies. The authors discovered that trisomy of chromosome 1q in particular is advantageous to cancer cells and phenocopies the loss of tumor suppressor *TP53* signaling. Tumors with this aneuploidy are sensitive to compounds activated by an enzyme encoded on chromosome 1q, suggesting a potential therapeutic approach. —YN

Science, adg4521, this issue p. 848

CELL BIOLOGY

How midnolin catches its substrates

Eukaryotic cells contain a macromolecular protease called the proteasome that degrades proteins modified by ubiquitin. The proteasome can also degrade proteins that are not ubiquitinated, but how this occurs mechanistically has remained mysterious. Gu *et al*. identified midnolin, an inducible protein that localizes within the nucleus to promote the proteasomal degradation of numerous transcriptional regulators independently of ubiquitination (see the Perspective by Schilling and Weber-Ban). Midnolin stably associates with the proteasome and uses a structural domain that incorporates a free β strand to "catch" substrates for destruction. Thus, the midnolin-proteasome pathway bypasses the canonical ubiquitination system to achieve selective degradation of many nuclear proteins. —SMH

Science, adh5021, this issue p. 849; see also adj8230, p. 834

STRUCTURAL BIOLOGY

Chromatin assembly by CAF-1

Chromosomal duplication involves the replication of DNA and the assembly of chromatin onto the replicated DNA. Chromatin assembly factor-1 (CAF-1) is a heterotrimeric protein complex responsible for the deposition of newly synthesized histones H3 and H4 onto DNA in replication-coupled nucleosome assembly. Liu *et al*. revealed that in the presence of histones H3 and H4, the two otherwise flexibly tethered globular domains of CAF-1 attach to opposite ends of an elongated H3-H4 heterodimer, and a flexible loop of the p150 subunit becomes ordered to secure H3-H4 binding. Formation of an H3-H4 tetramer is a prerequisite for nucleosome assembly, and DNA induces the juxtaposition of two CAF-1–H3-H4 complexes poised for H3-H4 tetramer formation. CAF-1 promotes right-handed DNA wrapping of H3-H4 tetramers in vitro, in contrast to the left-handed DNA wrapping in the nucleosome. —DJ

Science, add8673, this issue p. 850

SENSATION

Intimate touch

It is well known that the genitals are unusually sensitive, and that genital touch is crucial for mating and associated pleasure, but the underlying basis is not completely understood. By studying mice and humans with a rare inherited mechanosensory syndrome, Lam *et al*. identified a mechanism involving the mechanoreceptor PIEZO2 that is responsible for determining genital sensitivity (see the Perspective by George and Abraira). Their results highlight the importance of touch for driving physiological responses needed for sexual function. The identification of PIEZO2 and a specific type of touch neuron as key mediators might help in the development of therapeutic approaches for both hypo- and hypersensitivity that interfere with the enjoyment of sex. —MMa

Science, adg0144, this issue p. 906; see also adj8674, p. 832

CARBON OFFSETS

A good idea in principle

Reducing emissions from deforestation and forest degradation (REDD) projects are intended to decrease carbon emissions from forests to offset other carbon emissions and are often claimed as credits to be used in calculating carbon emission budgets. West *et al*. compared the actual effects of these projects with measurable baseline values and found that most of them have not reduced deforestation significantly, and those that did had benefits substantially lower than claimed (see the Perspective by Jones and Lewis). Thus, most REDD projects are less beneficial than is often claimed. —HJS

Science, ade3535, this issue p. 873; see also adj6951, p. 830

SOLAR PHYSICS

Numerous small-scale jets on the Sun

Plasma is constantly streaming away from the Sun, forming the solar wind. A likely source of this plasma is coronal holes, regions of the Sun's corona with magnetic field lines that open outward. Chitta *et al*. observed a coronal hole in the extreme ultraviolet using the Solar Orbiter spacecraft and identifed several types of small-scale jets within it (see the Perspective by Ugarte-Urra and Wang). Large numbers of jets occurred during the observation, but each one lasted only a few dozen seconds. The authors calculated that the jets provide enough energy and plasma to supply a large fraction of the solar wind, at least during quiet periods. —KTS

Science, ade5801, this issue p. 867; see also adj8002, p. 833

THYMUS

Aire not so apparent

The T cell repertoire is formed from a series of intricate steps during which autoreactive thymocytes are eliminated. Glucocorticoids, produced locally, can promote thymocyte survival, counteracting the negative selection facilitated by the expression of tissue-restricted antigens in thymic medullary epithelial cells that is induced by the gene *autoimmune regulator* (*Aire*). Using genetically engineered mice, Taves *et al.* found that medullary, but not cortical, epithelial cells showed enriched expression of genes encoding enzymes required for all steps of glucocorticoid biosynthesis, including Cyp11b1, and that their expression and the abundance of glucocorticoids in the thymus were dependent on Aire. Although the importance of Aire in limiting autoimmunity by regulating negative selection is established, Aire-dependent expression of enzymes involved in lipid and steroid biosynthesis may modulate the process to refine T cell diversity. —SHR

Sci. Immunol. (2023) 10.1126/sciimmunol.abo7975

ION CHANNELS

A multitude of ways to activate TPCs

The second messenger nicotinic acid adenine dinucleotide phosphate (NAADP) indirectly stimulates the opening of two-pore channels (TPCs) on endosomes and lysosomes, thus triggering the release of calcium ions from these acidic organelles. Gunaratne *et al.* found that the NAADP-binding

proteins JPT2 and LSM12 were both required in a nonredundant fashion for NAADP-evoked calcium release. In a related paper, Saito *et al*. found that a synthetic compound mimicked the effect of NAADP on TPC2-mediated calcium release in a manner distinct from that of NAADP-binding proteins, suggesting strategies for restoring TPC2-dependent calcium release in pathological contexts. —AMV

Sci. Signal. (2023)
10.1126/scisignal.adg0485,
10.1126/scisignal.adg0661

LIQUIDS REDEFINED

Liquids reveal solid-like properties

Recent advances in theoretical and experimental research have led to a reevaluation of our understanding of liquids, revealing that they have properties akin to solids. Trachenko points out that the conventional Maxwell model that describes liquids as purely viscous is not universal. The interaction between the liquid and the substrate is key, and under total wetting conditions, liquids demonstrate a solid-like response, including shear elasticity, which challenges the very definition of what a liquid should be. This shear elasticity allows for the propagation of thermal waves, revealing a thermoelastic behavior similar to that of solids. These insights point to new applications for controlled fluid transport in the health sciences. —LSB

Sci. Adv. (2023)
10.1126/sciadv.adh9024

SOCIAL SCIENCES

Policy can both cause and reduce poverty

Researchers studying poverty in the United States often focus on individualized risks, such as a lack of college education, or other potential causes of poverty, including cultural accounts. In a review of poverty research and a comparison of the US with other nations, Brady shows that the US has historically higher poverty than comparable nations, and this cannot be accounted for by a greater prevalence of expected risks or causes. Instead, policy decisions to penalize statuses thought of as risks (e.g., single motherhood), and to penalize them more severely than other nations, better explain the comparatively higher US poverty rate. However, policy change can reduce poverty, just as policy has generated it. —JSE

Sci. Adv. (2023)
10.1126/sciadv.adg1469

METABOLISM

Long noncoding RNA for slim mice

The gut microbiota of mice influences their host's metabolism by mechanisms mediated by small intestinal epithelial cells. During whole-transcriptome sequencing of epithelial cells, Wang *et al*. found several genes that encode long noncoding RNAs (lncRNAs). One of these genes, called *Snhg9*, varied in abundance in animals depending on the presence or absence of gut microbiota mediated by immune effectors. The authors found that the lncRNA interacted with and bound to cell cycle and apoptosis protein 2, which in turn bound to and inhibited sirtuin, a regulator of lipid metabolism. High levels of *Snhg9* seen in germ-free mice may be the reason that they stay slim, whereas animals with intact microbiota have lower levels of the lncRNA and are fatter. —CA

Science, ade0522, this issue p.851

HEART DISEASE

Aging gets on the heart's nerves

Changes in innervation of the heart can contribute to arrhythmias, and the risk of arrhythmias greatly increases with age. Wagner *et al*. uncovered a mechanism connecting these two phenomena. The authors studied young and aging mice and demonstrated that innervation of the heart decreases with age. Age-related accumulation of senescent cells promotes the release of semaphorin-3A, which reduces the density of neuronal axons in the heart. At the same time, aging is associated with a decrease in a microRNA that counteracts the effects of semaphorin-3A, further tipping the balance toward decreased innervation. These age-related losses in innervation could be reversed by treating the mice with senolytic drugs, suggesting a potential therapeutic approach. —YN

Science, ade4961, this issue p. 897

MATERIALS SCIENCE

Curved structures from flat sheets

Transforming a flat sheet into a curved object or vice versa usually leads to distortions such as wrinkles or cuts unless the Gaussian curvature is conserved, such as when turning a piece of paper into a cylinder or a cone. Drawing inspiration from bulliform cells in plants, which are believed to control the curvature in leaves to regulate water loss, Gao *et al*. designed thin panels with inflatable pneumatic cells that allow for pure in-plane contraction or bending or combinations of the two. Thus, the authors were able to transform a flat surface into a shape with a different Gaussian curvature, with implications for the design of soft robotics. —MSL

Science, adi2997 this issue p. 862

CANCER

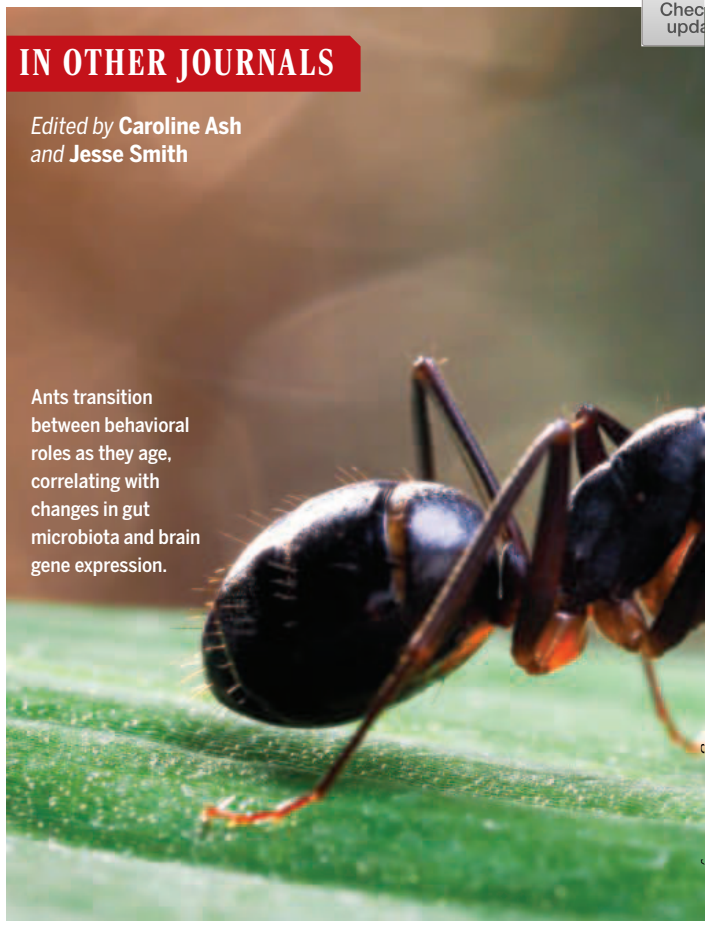
Improving immune checkpoint inhibitors

Although immune checkpoint inhibitors have been beneficial for many patients, 40% fail to respond. Bennion *et al*. evaluated Fcγ receptor IIB (FcγRIIB), an inhibitory receptor that has recently been found on activated $CD8^+$ T cells. They demonstrated that the Fc portion of the ligand antiprogrammed death 1 could bind to FcγRIIB on activated $CD8^+$ T cells in mice and reduce the efficacy of immunotherapy, which was abrogated by adding an antibody against FcγRIIB to the treatment. These results elucidate an unappreciated mechanism of FcγRIIB-mediated, cell-autonomous suppression, suggesting that altering the Fc portion of these antibodies could better improve efficacy across patients. —DLH

Sci. Transl. Med. (2023) 10.1126/scitranslmed.add1868

IN OTHER JOURNALS

Edited by **Caroline Ash** *and* **Jesse Smith**



Ants transition between behavioral roles as they age, correlating with changes in gut microbiota and brain gene expression.

PLANT PATHOLOGY

Synthetic resistance promoters

Plant pathogens cause yield reduction of crops worldwide. Different pathogens elicit immune responses through divergent plant signaling pathways, for example, by either activating salicylicacid (SA) or jasmonic acid (JA) biosynthesis. To enhance immunity to a broad range of pathogens, Li *et al.* designed synthetic promoter sequences that respond to both SA and JA. Chimeric promoters integrating both SA- and JA-responsive elements were activated by diverse plant pathogens. The promoters were used to drive the expression of an antimicrobial peptide conferring plants with enhanced resistance. This approach allows the induction of immune responses in response to otherwise antagonistic signaling pathways. —MRS

Plant Commun. (2023) 10.1016/j.xplc.2023.100596

PALEOECOLOGY

Foliage feeding inferred from fossils

Plants are hosts for hundreds of thousands of herbivorous arthropods. This diversity is thought to be maintained by the diversity of plants, with many herbivores specializing on certain

Damage preserved in leaf fossils provides evolutionary evidence for functional diversity among herbivorous arthropods.

BEHAVIOR

Social physiology

Colonies of social insects have sharply defined divisions of labor, from nursing and guarding to foraging and cleaning. Ant colonies thus provide convenient models for exploring the relationships between an animal's biology and its social environment. To disentangle the multiple contextual factors that add complication and redundancy to a social environment, Kay *et al.* conducted a multiomics study of the carpenter ant *Camponotus fellah.* The ants transition between behavioral roles as they age, which correlates with changes in microbiota, spatial environment, and other aspects of their biology. Machine learning applied to the omics data revealed that physiological measures (gut microbiota and brain gene expression) were more strongly correlated with social maturity than with behavior, age, or physical environment. These results imply that social interactions may have a direct effect on brain function. —CA

PLOS Biol. (2023) 10.1371/journal.pbio.3002203

host plants and many having adaptations to feed on different plant tissues. Using data from leaf fossils collected in central Europe, Iceland, and Norway, Albrecht *et al.* investigated which of these factors played a larger role in supporting herbivorous arthropod diversity over the past 60 million years. Their analysis showed that herbivore functional diversity during the Cenozoic was most closely tied to changes in the degree of host partitioning (multiple feeding types co-occurring on the same leaf) than to plant diversity. —BEL

Proc. Natl. Acad. Sci. U.S.A. (2023) 10.1073/pnas.2300514120

QUANTUM OPTICS

Synchronizing photon emission

Optical quantum networking and communication technologies rely on the ability to generate and propagate single photons. Although generating single photons is not a problem itself, having them be indistinguishable and doing so in a synchronized way remains a challenge. The generation process is often stochastic, making it difficult to engineer multiphoton states. Davidson *et al.* removed the stochasticity from the process by using two rubidium gas clouds. Cloud-one is used to generate pairs of single photons, and cloud-two is used as quantum memory to store a single photon. Cloud-one generates signal and idler photons. The signal photon is then stored in the memory until the idler photon triggers its release. Having a route to efficiently generate synchronized single-photon emission should allow for more complex (and more useful) multiphoton states to be engineered on demand simply by repeating the generation-storage-retrieval process. —ISO

Phys. Rev. Lett. (2023) 10.1103/PhysRevLett.131.033601

NEUROSCIENCE

Estrogen's effects on mood

Fluctuations in the levels of reproductive hormones, for example, during postpartum or perimenopausal periods, are associated with mood disruption in women. However, the neural circuits responsible for these changes are not clear. Tao *et al*. found that depressive-like behaviors caused by ovarian hormone withdrawal in mice were mediated by a reduction in activity of a group of estrogen receptor-1–expressing neurons in the medial preoptic area of the hypothalamus. These neurons project to the ventral tegmental area and the periaqueductal gray, where they enhance the release of dopamine and serotonin, respectively. Both dopamine and serotonin have positive effects on mood, so this finding provides a mechanistic link between estrogen levels and depression. —SAL

Nat. Neurosci. (2023) 10.1038/s41593-023-01397-2

GEOPHYSICS

Subducting seamounts

Seamounts dot the ocean floor and get subducted along with tectonic plates at subduction zones. Bangs *et al.* used three-dimensional seismic imaging to help understand how seamounts affect deformation processes in the subducting Hikurangi subduction zone of New Zealand. The authors found that one of the seamounts imaged created a bulge in the overriding plate, leaving a tube-shaped sediment wake. This structure seems to be tied to an aseismic deformation known as slow slip. Although seamounts are also thought to cause thrust faults to lock, these observations show how the diverse heterostructure created by seamounts also allows for other types of deformations. —BG

Nat. Geosci. (2023) 10.1038/s41561-023-01186-3

EDUCATION

High school guidance pays dividends

Effects of secondary school guidance counselors on educational attainment are of similar magnitude to teacher effects and are most impactful for lower-income and lower-achieving students. Mulhern analyzed data from 146 Massachusetts high schools that quasirandomly assigned 224,563 students to 613 counselors, who assisted students on things such as course selection and college applications. More effective counselors increased rates of high school graduation, college attendance and persistence, and bachelor's degree completion. Benefits were driven more by improved assistance and information for students than by the development of cognitive or noncognitive skills. —BW

Amer. Econ. Rev. (2023) https://www.aeaweb.org/articles?id=10.1257/aer.20200847

RESEARCH ARTICLE SUMMARY

STRUCTURAL BIOLOGY

Structural insights into histone binding and nucleosome assembly by chromatin assembly factor-1

Chao-Pei Liu*†, Zhenyu Yu†, Jun Xiong†, Jie Hu†, Aoqun Song†, Dongbo Ding†, Cong Yu, Na Yang, Mingzhu Wang, Juan Yu, Peini Hou, Kangning Zeng, Zhenyu Li, Zhuqiang Zhang, Xinzheng Zhang, Wei Li, Zhiguo Zhang, Bing Zhu*, Guohong Li*, Rui-Ming Xu*

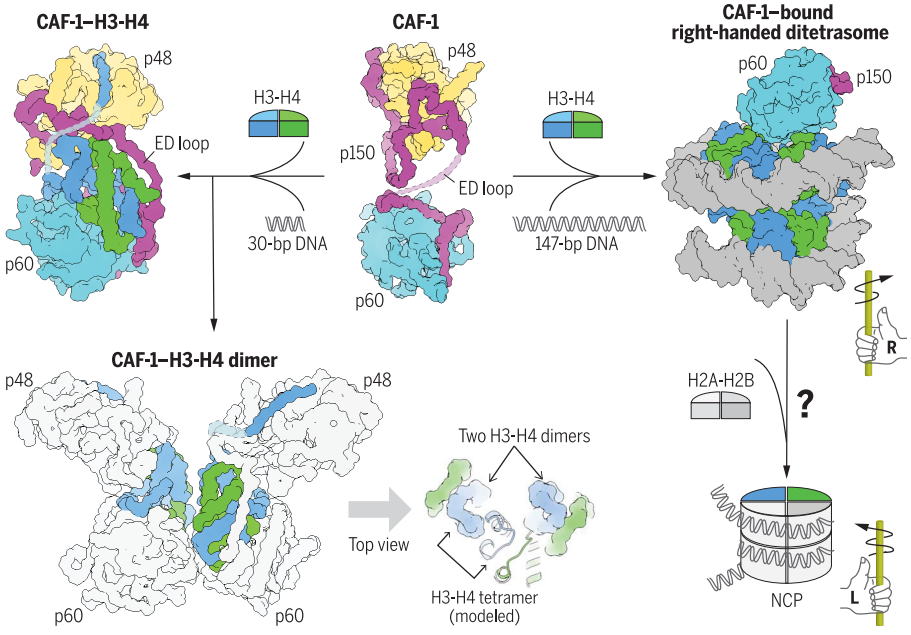
INTRODUCTION: Inheritance of chromatin structure during cell division is essential for genome integrity and epigenetic inheritance. DNA replication disrupts parental nucleosomes, and nascent nucleosomes must be assembled on replicated DNA. Half of the histones for the replication-coupled nucleosome assembly are of parental origin, and the other half are newly synthesized. Chromatin assembly factor-1 (CAF-1) is an evolutionarily conserved heterotrimeric protein complex responsible for the deposition of newly synthesized histones H3 and H4 onto DNA. A mechanistic understanding of the structure and function of CAF-1 is needed to fully comprehend the principles of chromatin inheritance.

RATIONALE: We determined the crystal structure of the core domain of human CAF-1 in the absence of histones and used cryo–electron microscopy (cryo-EM) to solve the structure of CAF-1 bound to histones H3 and H4. Structural findings were corroborated with in vitro supercoiling assays and in vivo nascent nucleosome mapping analyses. We also investigated the handedness of the nucleosome precursor assembled by CAF-1 using a single-molecule freely orbiting magnetic tweezer (FOMT) method.

RESULTS: The crystal structure of the CAF-1 core domain shows that, in the absence of histones, it adopts a bilobal structure, with p48 and p60 subunits freely tethered by the middle domain of the p150 subunit. Upon the binding of an H3-H4 heterodimer, p48 and p60 hug the opposite ends of the elongated H3-H4 heterodimer, whereas the negatively charged ED loop of p150 secures the binding by traversing the hunched positively charged surface of the H3-H4 dimer. The C-terminal portion of the ED loop plays a particularly important role in CAF-1's histone binding and nucleosome assembly activities, as shown by in vitro histone binding and plasmid supercoiling assays, as well as in vivo nucleosome assembly analysis. The observed histone binding mode ensures that only one H3-H4 heterodimer is bound by a CAF-1 complex because the joining of a second H3-H4 dimer in the manner of a H3-H4 tetramer would be sterically prohibited by p60, which suggests that formation of an H3-H4 tetramer is regulated. A p60 cancer mutation located at the interface with H4 is shown to have a detrimental effect in cell proliferation and cause a global change of gene expression. These effects likely reflect CAF-1's functions in chromatin accessibility and heterochromatin integrity.

We demonstrate that the addition of short DNA oligomers promotes the dimerization of the CAF-1–H3-H4 complex. A cryo-EM structure of a 2:2 CAF-1–H3-H4 complex shows that the two H3-H4 heterodimers are placed nearby, with the H3 dimerization interface poised for interaction but not quite in the geometric configuration of an H3-H4 tetramer. Given that longer DNA fragments wrap H3-H4 as tetramers, the length of DNA may be a key factor in the CAF-1–mediated formation of H3-H4 tetramers. With a 147–base pair (bp) nucleosome-positioning Widom 601 DNA, we found that a right-handed CAF-1–bound ditetrasome was assembled through salt dialysis. This discovery was confirmed by single-molecule FOMT at a physiological salt concentration, indicating a possible unusual nucleosome precursor in chromatin replication.

CONCLUSION: Our study reveals the histone binding mode of CAF-1. It also elucidates the role of DNA in the dimerization of the CAF-1–H3-H4 complex and the assembly of H3-H4 tetramers and suggests the involvement of a right-handed nucleosome precursor in replication-coupled nucleosome assembly. ■



Structures of CAF-1 with and without histones H3 and H4 bound. (Top middle) A 3.5-Å crystal structure of the core domain of CAF-1. Distinct subunits are labeled and color coded as indicated. The light pink segment with a dashed outline indicates disordered ED loop of p150. (Top left) A 3.5-Å cryo-EM structure of CAF-1 bound to a heterodimer of H3 (blue) and H4 (green). The light blue segment with a dashed outline indicates the disordered H3 loop that connects a short N-terminal segment and the body of H3. (Bottom left) A 4.6-Å cryo-EM map of a 2:2 CAF-1–H3-H4 complex in which CAF-1 subunits are colored gray. The graphic to the right of the arrow is a top view of the 2:2 dimer with CAF-1 omitted for viewing clarity. It shows the positioning of two H3-H4 heterodimers, shown in densities, in the 2:2 complex, compared with those in an H3-H4 tetramer, which is modeled by placing a fictitious H3-H4 dimer (ribbon representation) next to an observed one. (Top right) A 3.8-Å cryo-EM map of the CAF-1–bound right-handed ditetrasome compared with the left-handed DNA wrapping of a nucleosome core particle (NCP) schematically drawn at the bottom right corner.

The list of author affiliations is available in the full article online.
*Corresponding author. Email: liuchaopei@ibp.ac.cn (C.-P.L.); zhubing@ibp.ac.cn (B.Z.); liguohong@ibp.ac.cn (G.L.); rmxu@ibp.ac.cn (R.-M.X.)
†These authors contributed equally to this work.
Cite this article as C.-P. Liu *et al.*, *Science* **381**, eadd8673 (2023). DOI: 10.1126/science.add8673

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.add8673

RESEARCH ARTICLE

STRUCTURAL BIOLOGY

Structural insights into histone binding and nucleosome assembly by chromatin assembly factor-1

Chao-Pei Liu[1]*†, Zhenyu Yu[1]†, Jun Xiong[1]†, Jie Hu[1]†, Aoqun Song[1]†, Dongbo Ding[1]†, Cong Yu[1,2], Na Yang[3], Mingzhu Wang[4], Juan Yu[1], Peini Hou[1], Kangning Zeng[1,5], Zhenyu Li[6], Zhuqiang Zhang[1], Xinzheng Zhang[1,5], Wei Li[1,5], Zhiguo Zhang[7], Bing Zhu[1,5,8]*, Guohong Li[1,5]*, Rui-Ming Xu[1,2,5]*

Chromatin inheritance entails de novo nucleosome assembly after DNA replication by chromatin assembly factor-1 (CAF-1). Yet direct knowledge about CAF-1's histone binding mode and nucleosome assembly process is lacking. In this work, we report the crystal structure of human CAF-1 in the absence of histones and the cryo–electron microscopy structure of CAF-1 in complex with histones H3 and H4. One histone H3-H4 heterodimer is bound by one CAF-1 complex mainly through the p60 subunit and the acidic domain of the p150 subunit. We also observed a dimeric CAF-1–H3-H4 supercomplex in which two H3-H4 heterodimers are poised for tetramer assembly and discovered that CAF-1 facilitates right-handed DNA wrapping of H3-H4 tetramers. These findings signify the involvement of DNA in H3-H4 tetramer formation and suggest a right-handed nucleosome precursor in chromatin replication.

The basic unit of chromatin is the nucleosome core particle (NCP), which consists of ~147 base pairs (bp) of DNA wrapped around an octamer of histones H3, H4, H2A, and H2B in a left-handed form (*1*). During DNA replication, nucleosomes are first disassembled to allow the progression of the replication fork and then reassembled to package newly synthesized DNA into chromatin (*2*, *3*). Half of the histones for nucleosome assembly are of parental origin, and they are segregated onto two nascent DNA duplexes, mostly in the form of H3-H4 tetramers (*4*–*6*). The other half of the histones are newly synthesized in S phase of the cell cycle and are deposited onto DNA in a stepwise process known as de novo assembly, in which newly synthesized histones H3 and H4 are first deposited onto DNA to form tetrasomes, followed by the addition of two histone H2A-H2B dimers to form the NCP. The evolutionarily conserved chromatin assembly factor-1 (CAF-1) complex is responsible for DNA replication–coupled nucleosome assembly, which is essential for genome integrity and epigenetic inheritance (*7*–*9*). Despite much progress, further breakthroughs in structure and function studies of CAF-1, including its histone binding mode, are needed to fully comprehend the mechanism by which chromatin is faithfully inherited.

CAF-1 was discovered based on its ability to promote de novo nucleosome assembly during in vitro replication of SV40 DNA templates (*7*). It binds newly synthesized H3.1-H4 histones and facilitates their deposition at the replication fork during S phase in cooperation with histone H3-H4 chaperone ASF1 and the replication processivity factor proliferating cell nuclear antigen (PCNA) (*10*–*13*). Human CAF-1 is essential for replication-coupled chromatin assembly in vivo (*14*), but disruption of the corresponding yeast complex does not obviously affect cell growth (*15*, *16*). This is likely due to the presence of alternative chromatin assembly pathways in yeast (*12*, *13*, *17*–*19*). Additionally, CAF-1 plays important roles in chromatin repair after ultraviolet light–induced DNA damage (*20*, *21*) and in cell fate determination and maintenance (*2*, *22*).

Human CAF-1 is composed of p150 (CHAF1A), p60 (CHAF1B), and p48 (RBBP4/RbAp48) subunits (*9*, *23*). p150 contains various motifs for interacting with the replication machinery and chromatin factors, including the DNA-binding KER and WHD domains and the highly acidic ED region that is implicated in histone binding (*21*, *24*, *25*) (Fig. 1A). The N-terminal region of p150 was shown to be dispensable, whereas the middle and the C-terminal regions are important for binding histones and p60, respectively (*9*). Both the p60 and p48 subunits are WD40-repeat proteins that are known to form a β-propeller scaffold (Fig. 1A). p60 plays important roles in DNA synthesis and repair and participates in the progression of various malignancies (*26*–*28*). It contains a C-terminal ASF1-interacting motif (B domain) that may facilitate the transfer of H3-H4 to CAF-1 (*29*). Importantly, p60 directly binds H3-H4 and is required for CAF-1's nucleosome assembly activity (*9*, *30*). p48 is a versatile histone H3-H4 chaperone found in numerous histone-modifying enzymes and nucleosome-remodeling complexes (*23*, *31*). It is conceivable that it may couple CAF-1's function in nucleosome assembly with the installation of histone modifications (*32*, *33*).

Until now, inadequate knowledge about the structure of CAF-1 severely limited mechanistic understandings of de novo nucleosome assembly (*30*, *34*–*36*). In this work, we have determined the crystal structure of the human CAF-1 core complex and cryo–electron microscopy (cryo-EM) structures of CAF-1 bound to H3-H4. Interestingly, we also observed a complex of CAF-1 bound to a right-handed H3-H4 ditetrasome, implicating an unexpected step in nucleosome assembly.

Results

Overall structure of the CAF-1 complex

We first tried to crystallize the CAF-1 complex with full-length p150 or its truncation variants p150M, which contains the middle domain encompassing the ED region and its N- and C-terminal flanking regions (NTR and CTR, respectively), and p150L, which has a longer C-terminal extension (Fig. 1A). These complexes were coexpressed in insect cells and purified to apparent homogeneity (Fig. 1B), but only the complex with p150M, p60 with a C-terminal 140-residue deletion (p60ΔC), and full-length p48, termed CAF1-MC was crystallized. Three structures from distinct crystal forms ($C222_1$, C2, and $P2_1$) at 3.4- to 3.6-Å resolutions were solved (fig. S1A and table S1), and they all have a 1:1:1 subunit stoichiometry. p150M exhibits an extended conformation, with its ordered portions of NTR (residues 463 to 606) and CTR (residues 658 to 713) binding p48 and p60, respectively (Fig. 1C). An ~50-residue region connecting the NTR and CTR is disordered, and this results in an ambiguity in distinguishing two crystallographically equivalent models (Fig. 1C and fig. S1B). In model 1, the p48 and p150 NTR globular module makes little contact with the p60-p150 CTR module (Fig. 1C); in model 2, the two modules are close together, and an NTR segment (residues 524 to 542) interacts with p60 (fig. S1, B and C). Glutathione *S*-transferase (GST)–pulldown results show that the p150 NTR and CTR fragments only pulled down p48 and p60 separately (fig. S1D), thus disfavoring the model 2 packing. Thus, we used model 1 for further analysis.

[1]National Laboratory of Biomacromolecules and Key Laboratory of Epigenetic Regulation and Intervention, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China. [2]Key Laboratory of Systems Health Science of Zhejiang Province, School of Life Science, Hangzhou Institute for Advanced Study, University of Chinese Academy of Sciences, Hangzhou, Zhejiang 310024, China. [3]State Key Laboratory of Medicinal Chemical Biology, College of Pharmacy and Key Laboratory of Medical Data Analysis and Statistical Research of Tianjin, Nankai University, Tianjin 300353, China. [4]Institutes of Physical Science and Information Technology, Anhui University, Hefei, Anhui 230601, China. [5]School of Life Sciences, University of Chinese Academy of Sciences, Beijing 100049, China. [6]School of Physical Sciences, University of Chinese Academy of Sciences, Beijing 100049, China. [7]Institute for Cancer Genetics, Department of Pediatrics and Department of Genetics and Development, Columbia University Irving Medical Center, New York, NY 10032, USA. [8]New Cornerstone Science Laboratory, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China.
***Corresponding author. Email: liuchaopei@ibp.ac.cn (C.-P.L.); zhubing@ibp.ac.cn (B.Z.); liguohong@ibp.ac.cn (G.L.); rmxu@ibp.ac.cn (R.-M.X.)**
†These authors contributed equally to this work.

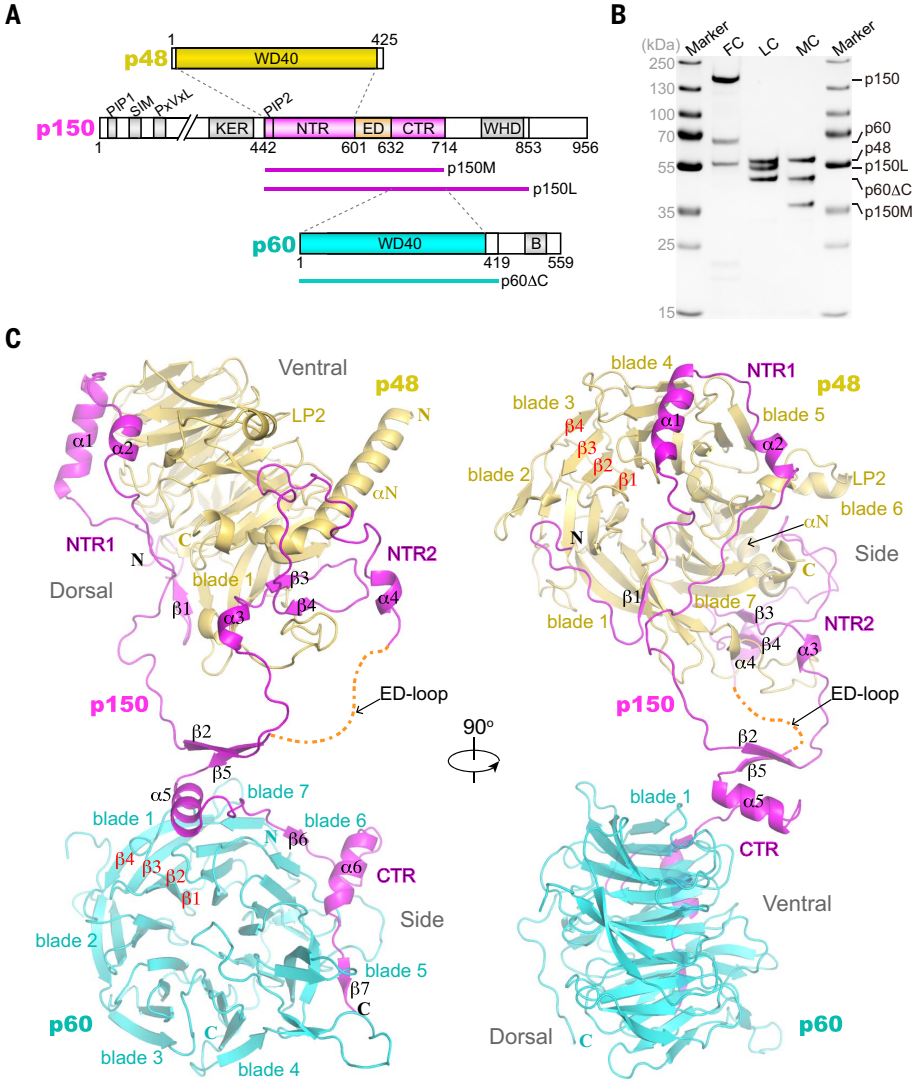


Fig. 1. Crystal structure of CAF-1. (**A**) Domain structures of CAF-1 subunits. The core domains of p48, p150, and p60 are colored amber, magenta, and cyan, respectively. Line segments that are colored the same as the protein domains indicate protein fragments used in this study. Dashed lines indicate protein regions involved in pairwise interactions. (**B**) Coomassie blue–stained SDS–polyacrylamide gel electrophoresis (SDS-PAGE) analysis of purified recombinant CAF-1 complexes. Displayed are CAF1-FC; CAF1-LC, containing p150L, p60ΔC, and p48; and CAF1-MC, containing p150M, p60ΔC, and p48. (**C**) Crystal structure of CAF-1. Different subunits are colored according to the scheme in (A), and the orange dashed line denotes the disordered ED domain of p150. The seven blades of the p48 and p60 β propellers, from the N- to C-terminal direction, are numbered from 1 to 7. Four strands in each blade are numbered from β1 to β4, as illustrated in blade 1 of p60 and blade 3 of p48.

The p150 NTR is anchored to two disjoint surface patches on p48. One p150 binding region occupies a large surface area on the wider base of the p48 β propeller, termed the dorsal face (Fig. 1C). The N-terminal portion of p150 NTR, which encompasses residues ~463 to 533 and is designated NTR1, is composed of a short β strand (β1), two α helices (α1 and α2), and irregular loop segments. NTR1 does not have a compact core, and it snakes across the dorsal surface of p48 mainly through hydrophobic interactions. The second p48-interacting region of p150 NTR, called NTR2, approximately spans residues 548 to 606 and is connected to NTR1 by an ~15-residue linker that is not interacting with either p48 or p60. NTR2 contains two short α helices (α3 and α4) and a pair of short β strands (β3 and β4). NTR2 interacts with p48 through parallel β pairing between its β3 and the outmost strand of p48's blade 7, as well as through its β3-β4 and β4-α4 loops, which are arranged like a pair of claws, gripping on the C-terminal portion of the αN helix of p48 (Fig. 1C). The combined interactions through NTR1 and NTR2 stably anchor the NTR of p150 onto p48.

p60 is a seven-bladed β propeller characterized by a long β-hairpin insertion between β4 of blade 4 and β1 of blade 5 on the ventral surface, a long β-loop insertion between β1 and β2 of blade 6, and a long C-terminal tail that traverses the dorsal surface of the β propeller (Fig. 1C). The CTR of p150 interacts with p60 through two α helices (α5 and α6) and two β strands (β6 and β7), whereas its β5 makes an intramolecular β pairing with β2 of NTR. α5 of p150 CTR interacts with p60 mainly through polar interactions with three loop segments of p60 that line the ventral edge of the β propeller. The β6-α6-β7 cassette of p150 CTR makes extensive interactions with blades 5 and 6 of the p60 β propeller, with β6 and β7 making antiparallel β pairings with β4 strands of blade 6 and blade 5 of p60, respectively, whereas α6 and its flanking loop segments primarily interact with blade 6 (Fig. 1C).

The bilobal organization of the CAF-1 core complex indicates considerable conformational plasticity between the two globular modules, suggesting that dynamic relative positioning of p48 and p60 may be intrinsic to CAF-1's histone binding and nucleosome assembly activities.

Structure of CAF-1 bound to histone H3-H4

Extensive cocrystallization trials of H3-H4 with full-length and truncation mutants of CAF-1, which were tested by GST-pulldown and DNA supercoiling assays (fig. S2), were unfruitful. We then resorted to single-particle cryo-EM for structure determination and succeeded in determining the structures of the full-length CAF-1 complex (CAF1-FC) and CAF1-LC, a CAF-1 complex consisting of p150L that includes the WHD domain, full-length p48, and p60ΔC (Fig. 1B and fig. S2B), bound to a heterodimer of histones H3-H4 at 3.6- and 3.8-Å resolutions, respectively (figs. S3 and S4 and tables S2 and S3). Representative EM map sections of various parts of the 3.8-Å structure are shown in fig. S5. The two structures are highly similar in ordered regions and their spatial arrangements; their EM maps can be superimposed with a correlation coefficient of 0.976, and the datasets can be combined to yield a map of 3.5 Å (fig. S4 and table S3). Henceforth, instead of specifying the structures determined using the full-length CAF-1 complex (CAF1-FC) or the longer fragment CAF-1 complex (CAF1-LC), we will collectively refer to the two structures as CAF-1 for clarity of description, unless explicitly noted. In the structure, only the core region of p150, approximately spanning residues 491 to 713, and the WD40 domains of p60 (residues 1 to 392) and p48 (residues 11 to 411) are ordered (Fig. 2, A and B), generally corresponding to the CAF-1 core in the crystal structures. No densities for other parts of CAF-1 can be located, indicating

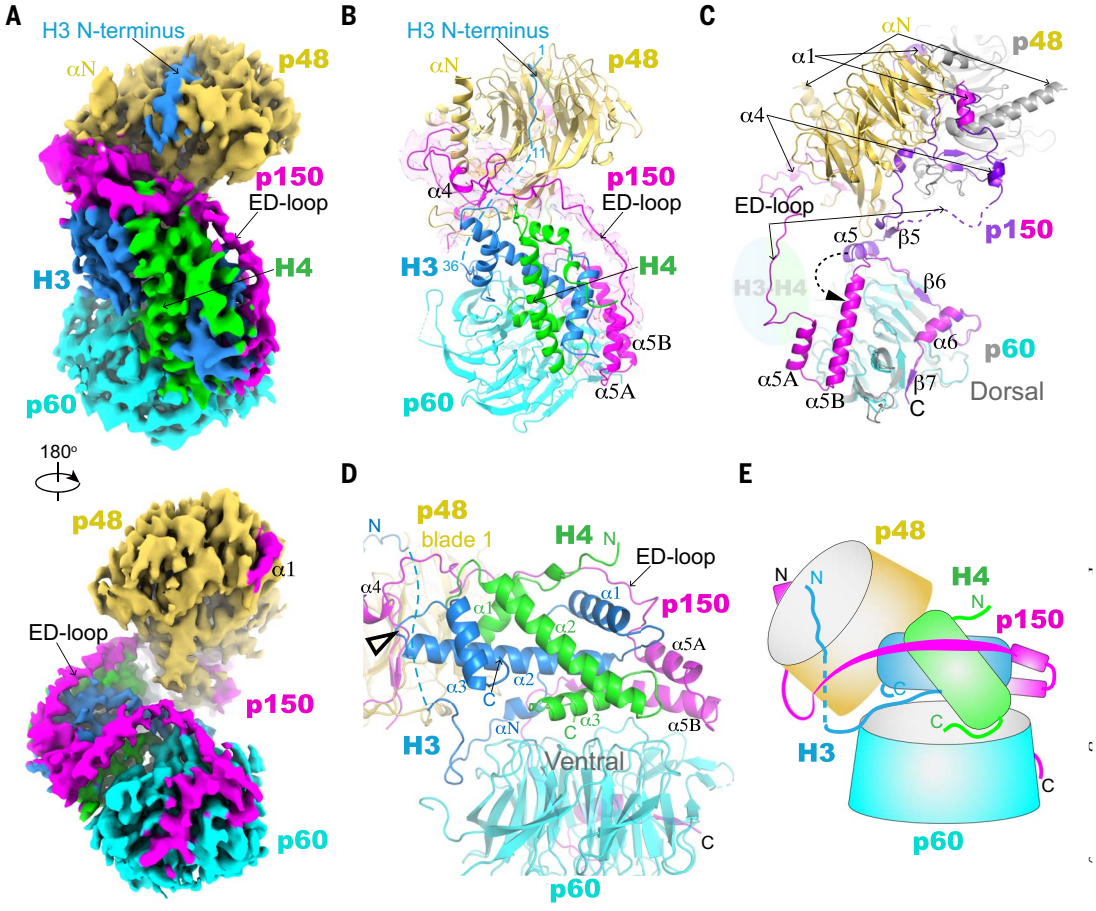


Fig. 2. Cryo-EM structure of CAF-1 bound to H3-H4. (**A**) Front (top) and back (bottom) views of the density map, which is sharpened with a B-factor of −144 Å^2 and contoured at 4.0σ. CAF-1 subunits are colored the same as in Fig. 1C, and H3 and H4 are shown in dark blue and green, respectively. (**B**) A cartoon model of the CAF-1–H3-H4 complex structure in the front view. p150 is highlighted with the superposition of a semitransparent surface representation of the density. Disordered internal loops are represented with dashed lines. (**C**) A back view of the superimposed apo crystal structure and the H3-H4–bound cryo-EM structure of CAF-1 through alignment of p60. The crystal structure is colored with p150 in periwinkle and p60 and p48 in gray; the cryo-EM structure is colored the same as in (B). Histones H3-H4 are represented by a pale bicolor ellipse for viewing clarity. Paired arrow lines indicate the relocation of the same structural elements of p150 and p48 in the two structures. Transformation of the β5-α5 unit of p150 in the crystal structure into part of α5B in the cryo-EM structure is indicated by a dashed arrow arc. See also movie S1. (**D**) An overview of the interaction between CAF-1 and histones H3-H4. The exposed histone H3-H4 tetramerization interface is indicated by the black triangle. (**E**) A schematic drawing depicting the overall features of H3-H4 binding by CAF-1.

that they are flexibly positioned relative to the CAF-1 core and that their interactions with H3-H4, if any, are dynamic and/or heterogeneous. There is also no density for a 30-bp DNA oligomer (DNA_{30}), which was added as a buffering reagent to alleviate histone precipitation at low-salt concentrations during cryo-EM sample preparation.

All three subunits of CAF-1 are involved in direct interaction with the H3-H4 heterodimer (Fig. 2, A and B). Compared with the crystal structure, the p48 and p60 modules showed little changes, but a large rearrangement of their relative positioning occurred (Fig. 2C and movie S1). Superimposing p60 from the two structures as a reference, p48 is rotated and located further away from p60 in the apo crystal structure than in the complex with H3-H4. For example, two p150 α4s from the two structures are located ~70 Å apart (Fig. 2C and movie S1). Through this substantial conformational change upon H3-H4 binding, the NTR1 region of p150 that interacted with p48 becomes largely disordered, with only a portion of α1 remaining visible. The linker between NTR1 and NTR2 is also disordered, but NTR2 binds p48 in the same manner as in the apo structure (Fig. 2B). The ~50-residue loop connecting NTR and CTR of p150, including the highly negatively charged ED loop, becomes ordered upon H3-H4 binding (Fig. 2A and fig. S5B). Within the CTR region of p150, the β6-α6-β7 cassette bound p60 essentially the same in the two structures (Fig. 2C). However, β5 is dissolved and joins α5 to form a longer α5B helix that orients in a direction opposite to the original α5 and rests on top of the ventral surface of p60, together with a newly ordered α5A that lies antiparallel to α5B (Fig. 2C and fig. S5B).

The ordered H3 regions include residues 1 to 11 and 36 to 134, whereas those of H4 include residues 21 to 101 (Fig. 2B and fig. S5B). With the exception of the first 11 residues of H3, which bind p48 in the same manner as they bind to other RbAp46/48 homologs (*33*), H3-H4 binds CAF-1 in a new fashion. H3-H4 is stably positioned above the ventral surface of the p60 propeller by means of extensive interactions involving H4's C-terminal part, from the penultimate turn of α2 (Tyr^{72}) onward (Fig. 2D). Furthermore, an H3 region encompassing residues 40 to 55, which includes a partially unfolded αN, wraps around a contiguous perimeter margin of the p60 β propeller spanning blades 6, 7, and 1. With p150's NTR and CTR firmly anchored on p48 and p60, respectively, the ED domain of p150 adopts an extended conformation and binds H3-H4 by traversing the positively charged convex surface of H3-H4 that is formed by the α1 helices of H3 and H4 (Fig. 2D). The newly formed α5A and α5B of p150 are positioned above the ventral surface of p60 and contact the α2-α3 loop of H4, the α1-α2 loop, and the partially unfolded αN region of H3. Intuitively, one may imagine that CAF-1 subunits house the binding of H3-H4 like a car seat (Fig. 2E), with the ventral surface of p60 as the seat, the long positively charged ED loop of p150 as the seatbelt, α5A and α5B of p150 as the guardrail, and the side of p48 as the headrest, against which the C-terminal end of H3's α2 and the α1-α2 loop of histone H4 rest.

Interaction between p150 and histones H3-H4

The ED loop of p150 (residues 601 to 632) interacts extensively with both histones H3 and H4 (Fig. 2D). The rest of the interaction mainly involves α5A and α5B, which are located immediately C-terminal to the ED loop. Two-thirds

of the ED-loop residues are negatively charged, and they interact with H3-H4 following a trail of positively charged residues located on the curved back of the H3-H4 heterodimer, which is mainly formed by the α1 helices of H3 and H4 that are normally involved in binding DNA in the nucleosome (Fig. 3A). The charge interaction between the ED loop of p150 and H3-H4 is somewhat reminiscent of that between the histone chaperone domain of MCM2 and H3-H4 (fig. S6A) (*37*–*39*), indicating a common strategy for safeguarding H3-H4 from promiscuous interaction with DNA during DNA replication. The vast overlapping binding area of CAF-1 and MCM2 on H3-H4 may protect their binding fidelity toward nascent and parental histone H3-H4s, respectively (*40*).

The main-chain model of the 32-residue, surface-exposed ED loop fits the continuous cryo-EM map well, but weaker side-chain densities limited precise identification of specific interactions between the ED loop and histones H3-H4 (fig. S5B). To circumvent this problem, we carried out bloc mutations on both the ED loop and histones to test their interactions. The N- and C-terminal halves of the ED loop, ED-N1 (residues 601 to 617) and ED-C1 (residues 618 to 632), and a longer C-terminal fragment, ED-C2 (residues 617 to 646) that includes α5A, were replaced with equal-length GS linkers, which are tandem penta-peptide repeats consisting of four contiguous glycine residues and a serine (Fig. 3A). Helix α5A interacts with the α1-α2 loop of H3 and the α2-α3 loop of H4 (Fig. 2D). GST-pulldown (Fig. 3B) and in vitro supercoiling (Fig. 3C and fig. S6B) assays showed that the ED-C1GS and ED-C2GS substitutions weakened the histone binding and nucleosome assembly abilities of CAF-1, to degrees comparable to that when the entire ED-loop was substituted (EDGS), whereas the ED-N1GS substitution displayed less appreciable effects. Amino acid substitution of four negatively charged residues in ED-C1 [denoted 4mut, with substitutions E628K (Glu^{628}→Lys), D629N, E630K, and D632G] displayed even weaker effects in the GST-pulldown and supercoiling analyses (Fig. 3, B and C). Furthermore, because ED-N1 mainly interacts with histone H4 and ED-C1 and ED-C2 with H3 (Fig. 3A), we probed reciprocal interactions using H3 and H4 mutants. Alanine substitutions of the ED loop–interacting residues Arg^{63}, Arg^{69}, and Arg^{83} of histone H3 (H3-3A) and Arg^{35}, Arg^{39}, and Lys^{44} of histone H4 (H4-3A) were generated, and the H3-3A and the combined H3-3A H4-3A (6A) mutant complexes exhibited weakened binding to CAF-1 compared with the wild-type (WT) H3-H4 complex, whereas the H4-3A complex only displayed subtle effects (fig. S6C). These results support the structural model of p150-histone interaction and highlight the importance of the interaction between the C-terminal portion of the ED loop and histone H3.

CAF-1's replication-coupled (RC) nucleosome assembly activity directly affects nucleosome occupancy behind the replication fork and is also required for optimal gene silencing and S phase progression (*14*–*16*, *41*). Therefore, we used two methods to analyze the cellular effects of relevant CAF-1 mutants: (i) a modified ReIN-Map (replication-intermediate nucleosome mapping) method combined with MINCE-seq [mapping in vivo nascent chromatin with 5-ethynyl-2′-deoxyuridine (EdU) sequencing] (*42*, *43*) and (ii) transcriptome analysis together with a cell proliferation assay. We first established p150 and p60 auxin-inducible degron (AID) HAP1 cell lines by CRISPR-Cas9–mediated knock-in, with which the AID-tagged p150 or p60 proteins could be degraded within hours after 5-phenyl-indole-3-acetic acid (5-Ph-IAA) treatment (*44*) (fig. S6D). We then introduced doxycycline-inducible exogenous p150 constructs into the p150-AID cell line by lentiviral transduction (fig. S6E). The cells cannot survive prolonged p150 depletion by treating with the auxin analog 5-Ph-IAA (6 days), and we found that the p150 ED-loop replacement mutants affect cell proliferation to degrees analogous to the levels to which they influence histone binding, with the 4mut substitution mutant showing the least growth defect (Fig. 3D). Analysis of RC nucleosome assembly was performed following a procedure schematically outlined in Fig. 3E. Analyses of EdU-labeled nucleosome occupancies surrounding transcription start sites (TSSs) and CTCF-binding sites showed again that the p150 ED-loop mutants differentially affect RC nucleosome assembly in a manner dependent on their histone binding abilities (Fig. 3F). By comparison, steady-state nucleosome occupancy is little affected (fig. S6F). Depletion of the CAF-1 subunits has been well documented to induce derepression of silenced genes, likely because of impairment of heterochromatin formation (*16*, *45*). Therefore, gene derepression analysis is a good surrogate for CAF-1 function in chromatin assembly. RNA sequencing (RNA-seq) analyses revealed that induced degradation of p150 leads to elevated expression of ~2500 genes, in contrast to the ~250 down-regulated genes, which is consistent with the functional role of CAF-1 in heterochromatin formation and silencing. Rescue with exogenous WT p150 restored ~80% of expression patterns of WT AID cells (fig. S6G). The C1GS, C2GS, and EDGS mutants showed obvious defects (~30 to 40%), whereas the 4mut substitution mutant functioned nearly as well as the WT p150 (fig. S6G).

p60-histone interaction and implication for oligomeric states of H3-H4

p60 interacts with H3 and H4 primarily through its ventral surface area. The central area of the ventral surface is enriched with negatively charged residues, together with several tyrosines and a phenylalanine. It hosts extensive interactions with H4 residues located at the C-terminal end of α2, the entire α3, and the following C-terminal tail (Figs. 2D and 4A). In particular, Tyr^{88} and Lys^{91}, which are located on α3 of H4, insert into the central basin of the ventral surface of p60 (Fig. 4A). Tyr^{88} is surrounded by hydrophobic residues Tyr^{133}, Tyr^{173}, and Phe^{236} of p60, whereas Lys^{91} is positioned within bonding distances from Asp^{86} and Asp^{131} of p60. Recurrent missense mutations of H4 residues Lys^{91}, Tyr^{98}, and His^{75}, which are all located on the interface with p60, are found in a neurodevelopmental syndrome (*46*), whereas an Asp-to-Asn substitution of Asp^{86} (D86N) of p60 is found in somatic cancer mutations tabulated in the Catalogue of Somatic Mutations in Cancer (COSMIC) database. GST-pulldown and supercoiling assays of the p60 D86N mutant complex of CAF-1 showed no obvious defects (fig. S7, A and B), likely because, owing to the extensive p60-H4 interaction, a single D86N change is not sufficient to yield detectable effects in vitro. Nevertheless, we examined the cellular effect of this mutation using the p60 AID cell line (figs. S6D and S7C). The p60 D86N mutant caused a severe growth defect (Fig. 4B), suggesting its functional importance. Induced degradation of p60 increased the expression of 681 genes (Fig. 4C). The differential effects in cell growth and gene expression between the WT and D86N p60 may be largely attributed to the amino acid difference, because WT and D86N p60, and also p150, were expressed at comparable levels (fig. S7C), notwithstanding the reduction of p150 upon p60 degradation in the absence of exogenous p60 (fig. S6D), as also previously reported (*47*, *48*). Of the up-regulated genes from p60 degradation, 87% coincide with those from p150 degradation, consistent with their shared function in the same protein complex. Even those genes not belonging to the common set changed in the same direction (fig. S7D). Finally, ReIN-Map analysis revealed no obvious change by the p60 D86N mutation (fig. S7E), suggesting possible involvement in chromatin accessibility and heterochromatin integrity (*16*, *45*, *47*, *49*, *50*).

Besides H4, a partially unwound αN of H3 lies in a negatively charged channel on the p60 surface, with H3 Tyr^{54} interacting with Tyr^{346}, Asp^{367}, and Tyr^{369} and H3 Arg^{49} contacting Glu^{7} and Tyr^{369} of p60 (Fig. 4D). The H3 loop segment spanning residues 40 to 48 binds a positively charged surface patch located at the edge of the ventral surface of p60. In this binding manner, the dimerization interface of histone H3 is fully exposed (Fig. 2D). However, binding to an H3-H4 tetramer is prohibited, owing to steric clashes between p60 and histone H4 from a modeled second heterodimer (Fig. 4E). These observations reveal a mechanism of safeguarding premature H3-H4 tetramer

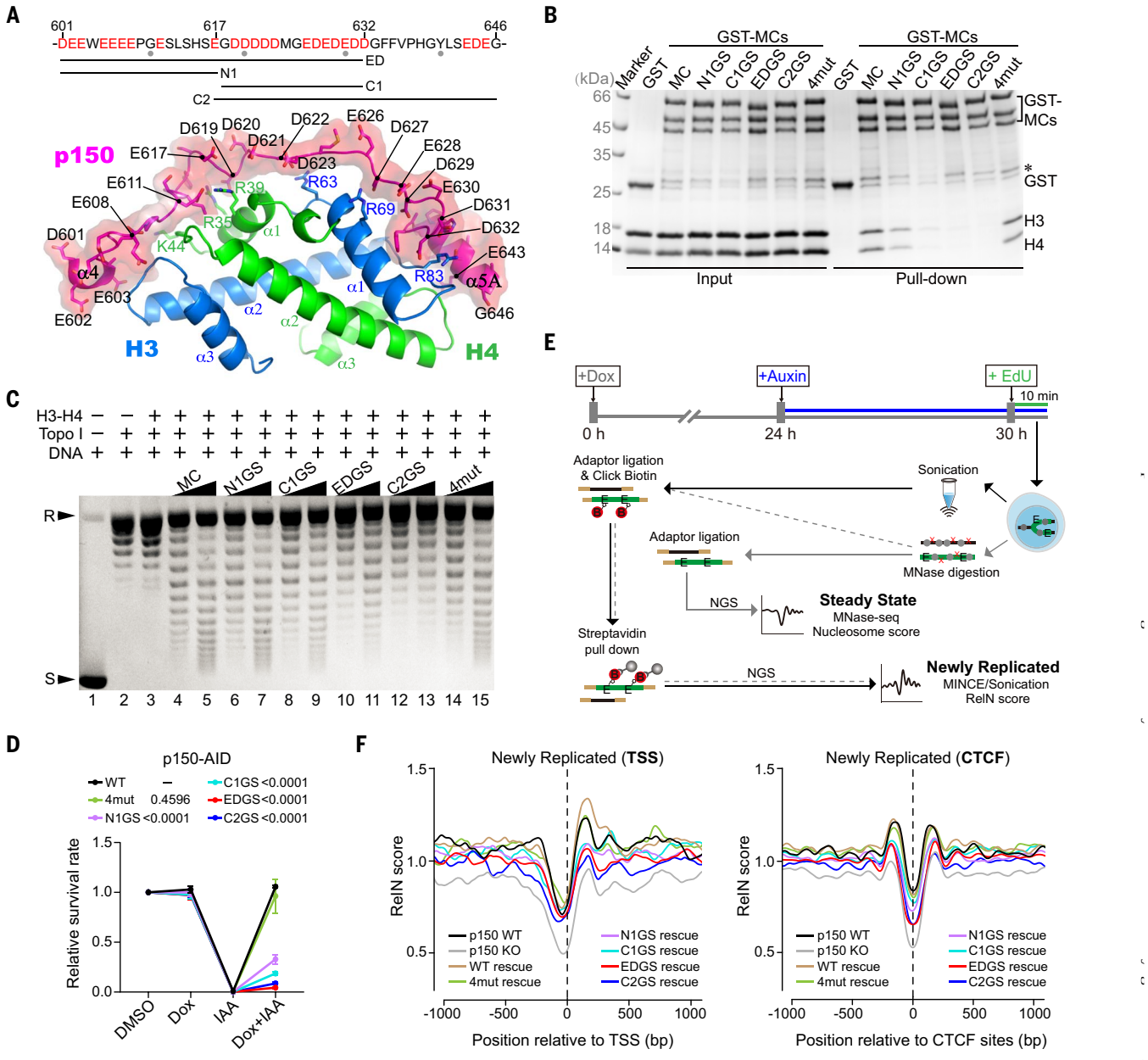


Fig. 3. Interaction between p150 and histones H3-H4. (**A**) Detailed view of the interaction between the ED domain of p150 and histones H3-H4. Shown at the top is the amino acid sequence of the ED domain, with acidic residues colored in red, and the designation of different fragments for functional analyses. Shown at the bottom is the interaction between key acidic residues in the ED domain of p150 and H3-H4. D, Asp; E, Glu; G, Gly; K, Lys; R, Arg. (**B**) Pulldown of H3-H4 with GST-tagged CAF1-MC complexes with an intact p150M or mutants substituting the ED-domain fragments in (A) with equal-length GS linkers; for example, N1GS and EDGS mutants substitute the N1 fragment and the entire ED domain with GS linkers, respectively. 4mut represents the E628K/D629N/E630K/D632G quadruple mutation in p150M. Pulldowns were performed at 1 M NaCl and analyzed by Coomassie blue–stained SDS-PAGE. The asterisk indicates an unidentified contaminant band. (**C**) Supercoiling assay of in vitro nucleosome assembly by CAF-1. ϕX174 DNA (lane 1) was treated with DNA topoisomerase I (Topo I, lane 2) and incubated with H3-H4 only (lane 3) or increasing amounts (0.1 and 0.2 μg) of WT or indicated mutants of CAF1-MC (lanes 4 to 15), together with 0.1 μg of H3-H4 and 0.1 μg of relaxed ϕX174 DNA. R and S at the left indicate the position of relaxed and supercoiled DNA, respectively. Input protein samples are shown in fig. S6B. (**D**) Cell proliferation activities of WT p150 or its mutants in p150-AID cells, which were first treated with doxycycline (Dox) to induce the expression of exogenous p150 constructs. 5-Ph-IAA was added 24 hours after Dox treatment. Cell viability was measured by CellTiter-Glo after 7 days of Dox treatment. Error bars represent SEM calculated from three biological replicates. Two-way analysis of variance (ANOVA) with Sídák test was used to calculate the *p* values shown at the top of the figure. (**E**) Schematic of the RelN-map and MNase-seq protocol. Cells were treated with Dox, Auxin, and EdU at the indicated time points, and chromatin was cross-linked and sheared either by MNase digestion or sonication. DNA of MNase-treated samples was extracted and ligated with next-generation sequencing (NGS) adaptors, and the MNase-seq signals reflect the steady-state nucleosome occupancy. For RelN-map, DNA of MNase-treated and corresponding sonication-sheared samples were extracted and ligated with NGS adaptors. EdU-labeled newly replicated regions were biotinylated and separated with streptavidin beads. The RelN score was derived by dividing the MINCE-seq signals by the Sonication-seq signals. (**F**) The RelN scores for regions around the TSSs (left) and the CTCF-binding sites (right) are shown. The CTCF chromatin immunoprecipitation sequencing (ChIP-seq) data are from the Gene Expression Omnibus (GEO; GSM4640493) (*81*).

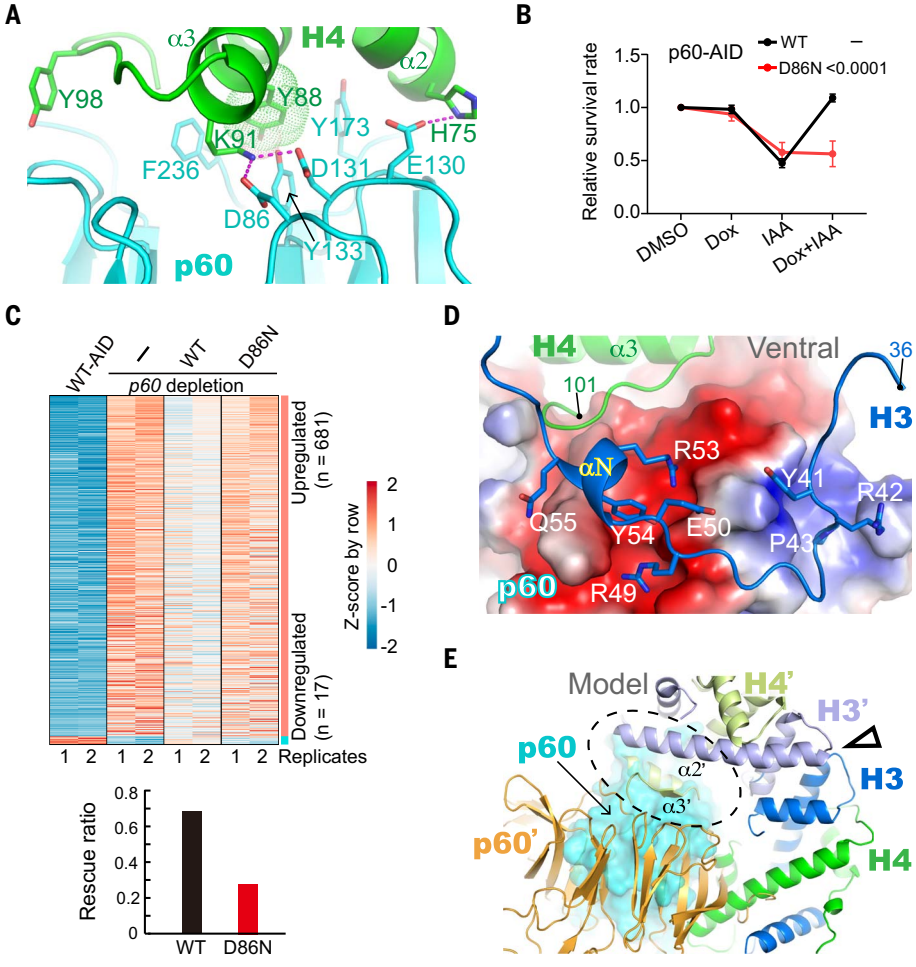


Fig. 4. Interactions between p60 and histone H3-H4. (**A**) Histone H4–p60 interaction at the center of the ventral surface of p60. Involved residues are shown in a stick model. Magenta dashed lines indicate potential hydrogen bonds. Tyr^{88} of H4, which is situated in the central basin of the ventral face of p60, is superimposed with a dots model. F, Phe; P, Pro; Q, Gln; Y, Tyr. (**B**) Cell proliferation activity of WT p60 and the D86N mutation assayed in the same was as in Fig. 3D. Error bars represent SEM from three biological replicates. (**C**) Heatmaps showing differentially expressed genes (DEGs) in p60-depleted and rescued cells versus WT-AID HAP1 cells. Two biological replicates were performed for each experiment. The bottom panel shows the rescue ratios of transcriptional changes, which is defined as $1 - [\#(A \cap B)]/[\#(A)]$, where A is the set of DEGs from p60 depletion and B is the set of DEGs from rescue expression of WT or D86N p60 constructs. The # function gives the number of genes in the set. (**D**) Interaction between histone H3 (dark blue) and p60. Histones H3-H4 are drawn in a cartoon model, with involved H3 residues depicted in a stick model. p60 is shown in a surface model colored according to electrostatic potential (negative, red; neutral, white; positive, blue) with the display range of −3 to +3 k_BT/e, where k_B is the Boltzmann constant and T and e are the temperature and electron charge in SI units, respectively. (**E**) Docking of a second CAF-1–H3-H4 complex (p60′, orange; H3′, light blue; H4′, light green) to model the formation of a H3-H4 tetramer between the H3′-H4′ heterodimer and the H3-H4 heterodimer from the first CAF-1–H3-H4 complex, which is colored as in Fig. 2B, results in steric clashes between H4′ and p60 (indicated with a dashed oval), and p60′ with p60 and H4. The black triangle indicates the H3-H3′ dimerization interface. p150 and p48 have been removed for viewing clarity.

formation by a histone H3-H4 chaperone, which differs from the mechanisms of maintaining H3-H4 in a dimeric form through the blocking of the dimerization interface by other H3-H4 chaperones (*51*).

Assembly of two nascent H3-H4 dimers into a tetramer has important implications for epigenetic inheritance, and the prevailing assumption is that H3-H4 heterodimers bound by two CAF-1 complexes will be spontaneously assembled into a tetramer (*24*). We observed a small percentage of particles with two CAF-1s bound to two H3-H4 heterodimers in the cryo-EM samples of both the LC and full-length complexes of CAF-1 bound to H3-H4 (figs. S3 and S4). The cryo-EM map calculated from combining the CAF-1 LC and FC datasets has an overall resolution of 4.6 Å, allowing reliable placement of two CAF-1–H3-H4 complexes (Fig. 5A and figs. S3 and S4). A SEC-MALS (size exclusion chromatography coupled to multiangle light scattering) analysis shows that the addition of 30-bp DNA, which was used in our cryo-EM sample preparation, promoted the dimerization of CAF-1–H3-H4 complexes, regardless of whether the MC, LC, or FC complex of CAF-1 was used (Fig. 5B and fig. S8). As in the cryo-EM reconstruction of the monomeric CAF-1–H3-H4 complex, we do not see the density of DNA here either. The structure shows that a pseudo twofold symmetry relates the two CAF-1–H3-H4 complexes. In this arrangement, the two p60s are juxtaposed, and a positively charged surface patch located on the side of the p60 propeller, which is formed by residues from the loop segments C-terminal to the β4 strands of blades 1 and 2 and the β3-β4 loop of blade 3, contacts the C-terminal portion of α2 of H4 from the other complex (Fig. 5C). The homodimerization interface on two H3s are facing each other but not oriented in the same way as in an H3-H4 tetramer (Fig. 5D). The two H3-H4 heterodimers in the structure appear to be "nearly" able to form a tetramer, but it is obstructed by the presence of p60. A remodeling of the observed p60-histone interaction, perhaps with the help of a longer DNA, would fulfill the assembly of an H3-H4 tetramer, which is believed to be the unit for CAF-1 deposition (*8*, *35*, *52*).

CAF-1 facilitates the assembly of a right-handed ditetrasome

Deposition of H3-H4 onto DNA by CAF-1 is believed to lead the formation of a tetrasome containing DNA and an H3-H4 tetramer. To gain further insights, we tried tetrasome assembly for structural characterization by salt dialysis in vitro. We initially dialyzed CAF1-LC, histone H3-H4, and a 147-bp Widom 601 DNA together from 2 M to 50 mM NaCl, and two major peaks, peak 1 and peak 2, appeared in cryo-EM sample preparation by GraFix (fig. S9A). From the peak 1 sample, the structures of a left-handed ditetrasome and a CAF-1–bound right-handed ditetrasome were determined at 3.5- and 3.8-Å resolutions, respectively (fig. S9, B to F; and table S2). The peak 2 sample predominantly shows a right-handed ditetrasome bound by two CAF-1 complexes, and the structure was determined at 5.6 Å (fig. S10 and table S2). The left-handed ditetrasome structure is nearly identical to the recently reported H3-H4 octasome structure (*53*); hence, we will not discuss it further here. We then repeated the cryo-EM procedures using the full-length CAF-1 complex and obtained similar results (fig. S11 and table S3). Cryo-EM reconstructions of a right-handed ditetrasome with one and two CAF-1 bound were obtained at 3.8- and 6.6-Å resolutions, respectively. Both

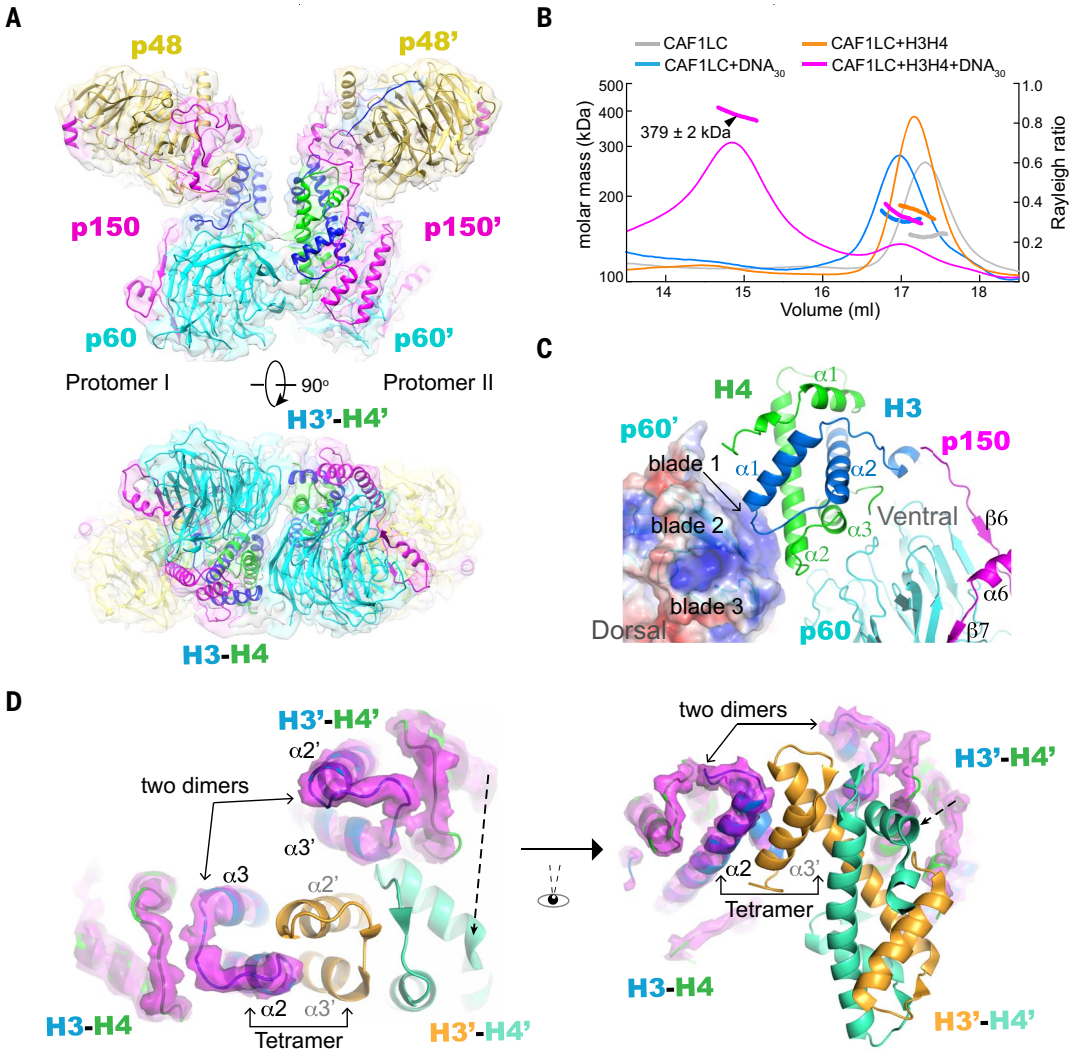


Fig. 5. A 2:2 complex of CAF-1 and H3-H4. (**A**) Two views of the 4.6-Å composite cryo-EM density map of the 2:2 complex contoured at 3.4σ superimposed with the structure model. The coloring scheme is the same as in Fig. 2A. (**B**) SEC-MALS analysis of CAF1-LC and its complex with H3-H4, in the presence and absence of 30-bp DNA. The SEC profile (thin lines) is superimposed with the fitted MALS results (short thick lines) according to the colors specified at the top of the figure. The 379 ± 2–kDa peak from the CAF1-LC–H3-H4–DNA_{30} sample indicates the formation of a 2:2 complex (calculated mass ~377 kDa). (**C**) Interactions between histone H3-H4 of one CAF-1–H3-H4 complex and p60′ of another complex, which is shown as the electrostatic potential surface (color range, −3 to +3 k_BT/e). H3-H4 binds a positively charged region spanning blades 1 to 3 of p60′. (**D**) Two views of the mismatched H3 dimerization areas, which are indicated by a pair of arrows labeled "two dimers," in the 2:2 complex. The two H3-H4 heterodimers from the cryo-EM structure are shown in a cartoon model superposed with a magenta semitransparent density map. For comparison with the face-to-face placement of two H3 dimerization surfaces in a tetramer, an additional H3-H4 heterodimer (cartoon model with H3′ in bright orange and H4′ in green cyan) is modeled onto one of the H3-H4 heterodimers in the 2:2 complex. The tetramer interface is indicated by a bidirectional arrow line labeled "tetramer."

the one–CAF-1–bound and the two–CAF-1–bound ditetrasome structures from the LC and FC complexes of CAF-1 are highly similar (fig. S11, E and F). Therefore, we will use the structures obtained from the CAF1-LC for the discussion that follows.

In the CAF-1–bound right-handed ditetrasome, a 108-bp DNA segment wraps around two overlapping H3-H4 tetramers ~1.5 times in a right-handed manner (Fig. 6A and fig. S12). Each H3-H4 tetramer binds DNA similarly to that in the nucleosome, except at the border of two tetramers, where the α2-α3 ends of H4s from two tetramers move in opposite directions laterally and closer longitudinally (Fig. 6B). As a result, the N-terminal ends of H3 α2 and H4 α3 of the abutting H3-H4 heterodimers form the interface between two H3-H4 tetramers (Fig. 6B). This dislocation at the interface of two H3-H4 tetramers is responsible for the right-handedness of the CAF-1–bound ditetrasome, which sees an opening of ~45 Å between the two gyres of DNA situated diametrically from the tetramer-tetramer interface (Fig. 6A). The peculiar H3-H4 tetramer interface is stabilized by the binding of p60 to the nearby C-terminal portion of H4 α2 by means of the β propeller's ventral surface edge spanning blades 2 and 3 (Fig. 6C). Additionally, p60 is stabilized on DNA through a positively charged surface area adjacent to its H4 binding region, spanning the contiguous ventral surface edge of blades 3 and 4. The DNA binding property of p60 was confirmed by an electrophoretic mobility shift assay (EMSA) (fig. S13A). Furthermore, the αN helices of two centrally located H3s belonging to different H3-H4 tetramers became ordered and interacted with DNA, H4 α1 of the same heterodimer, and H4 α2 from the nonadjacent heterodimer of the other H3-H4 tetramer (fig. S13B). The right-handed ditetrasome does not permit continuous addition

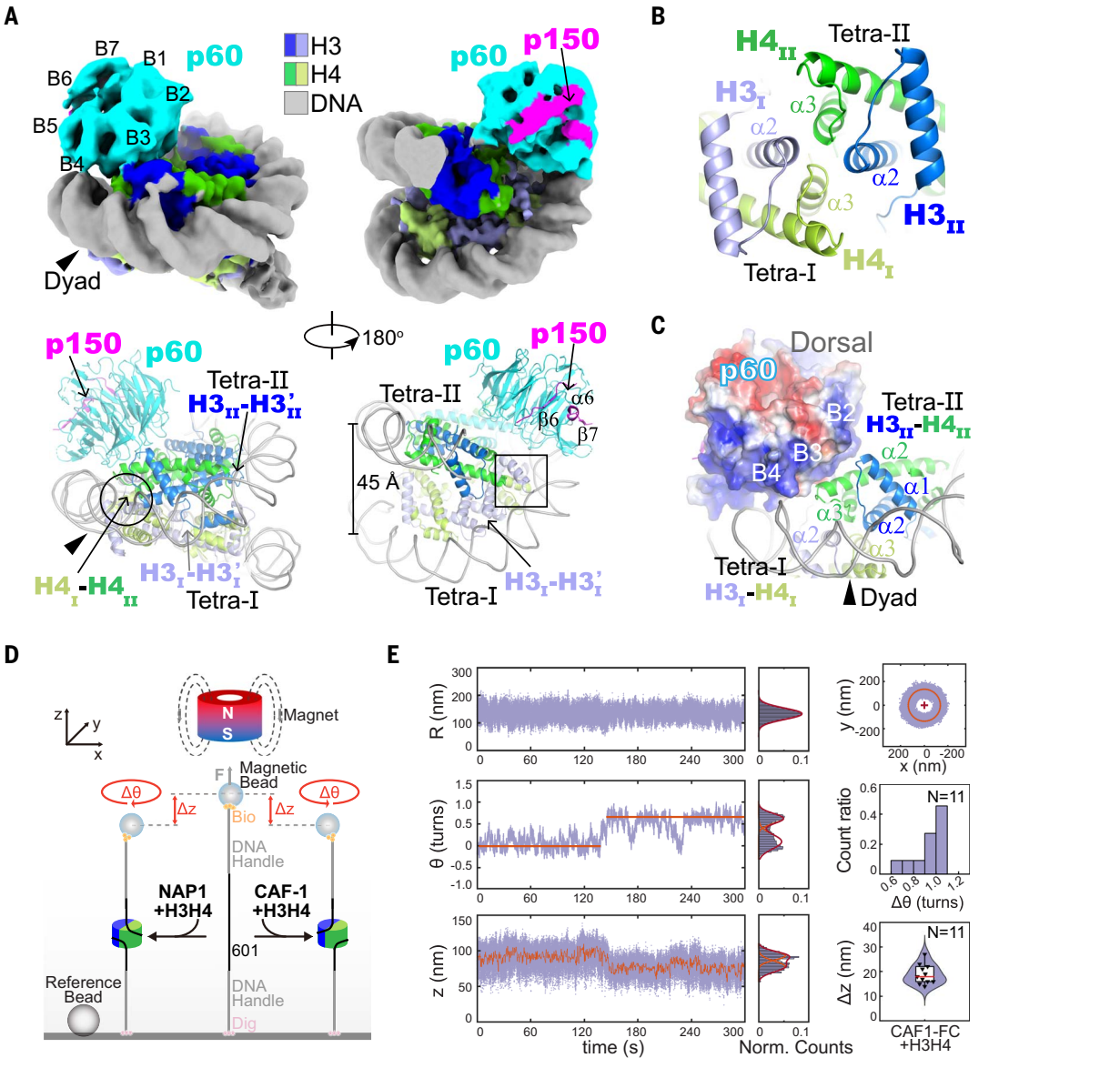


Fig. 6. CAF-1 induces right-handed DNA wrapping of H3-H4 in vitro. (**A**) Two views of the 3.8-Å resolution cryo-EM structure of the CAF-1–bound right-handed ditetrasome. Shown at the top is a cryo-EM density map contoured at 3.0σ. Regions for H3, H4, and DNA are colored according to the displayed color stripes, whereas that of p60 and p150 are colored as in Fig. 2A. Blades 1 to 7 of p60 are labeled as B1 to B7. A black arrowhead marks the dyad position at the interface between two H3-H4 tetramers, Tetra-I and Tetra-II. Shown at the bottom is a cartoon model of the structure. The two H3-H3 interfaces within the two tetramers and the H4 interface between Tetra-I and Tetra-II, which is highlighted in a circle, are labeled with arrows. The box in the right panel indicates the location of ordered αN of H3 in Tetra-I interacting with H4 of Tetra-II. (**B**) A close-up view of the inter-tetrasomal interface corresponding to the encircled region in (A). (**C**) Interaction of p60 with one H3-H4 heterodimer and DNA at the inter-tetrasomal interface. The electrostatic potential surface of p60 (−3 to +3 k_BT/e) is shown. The positively charged region spanning blades 2 to 4 contacts histones and DNA. (**D**) Schematic of the single-molecule FOMT for determining the handedness of DNA wrapping. The clockwise and counterclockwise circular arrows represent left- and right-handed turns from views looking down at the beads. (**E**) The top row shows the time course of radial fluctuation (left) based on the position of the bead (*x*, *y*) (right) and normalized counts of the left panel with Gaussian fit (middle) of a single DNA tether over a time span of 300 s covering an assembly event signified by simultaneous changes of bead rotation angle θ and extension length *z* shown below. The red circle in the right panel denotes the fitted overall circular motion of the bead. The middle row shows the time course of the bead rotation angle θ (left) and normalized distribution of θ changes with Gaussian fit (middle) of a single event. Mean values of each assembly step are indicated with red line segments. The histogram in the right panel shows the distribution of rotational changes (Δθ) from 11 independent tetrasome assembly events. The count ratio is defined as the number of times that a specific size rotation was observed divided by the 11 total events (*N* = 11). The bottom row shows the time course of the extension length (*z*) of a single DNA tether (left), with the red line showing *z* values averaged over 50-ms intervals, and normalized counts shown in the left panel with Gaussian fit (middle) of the same event as described above. The right panel shows the distribution of Δz values from the same 11 independent events as above displayed as a violin and box plot. The horizontal line represents the median, the height of the box extends the quartiles, and the whisker indicates the data range.

of H3-H4 tetramers because of steric clashes. Hence, discrete particles of ditetrasomes will be assembled on a long DNA. The extra density bound to p60 is consistent with being the β6-α6-β7 cassette of p150 CTR, which occupies the same location as p60 in the apo CAF-1 crystal structure and in the structure of CAF-1 bound to H3-H4; however, the rest of p150 and the entire p48 are missing (Fig. 6A). This observation suggests sizable conformational changes of CAF-1 upon deposition of H3-H4 onto DNA. A 5.6-Å resolution cryo-EM density map from the peak 2 sample shows the presence of two chunks of density clinging to the right-handed ditetrasome (fig. S10 and table S2), suggesting symmetric positioning of two p60-p150 units on both sides of the right-handed ditetrasome surface (fig. S10E). Finally, another distinct class of particles from peak 2 has a bigger lump of density bound to one side of a right-handed ditetrasome, resembling the presence of both p48 and p60 WD40 domains, but the low resolution of the map prevents further details from being elucidated (fig. S10C).

A nucleosome assembly intermediate, termed chromatin precursor, that consists of CAF-1, replicated DNA, and H3-H4 was previously predicted (*8*). It is surprising that we find a CAF-1–bound right-handed DNA ditetrasome in our in vitro assembly by salt dialysis. To exclude the possibility of an artifact from salt dialysis, we used a single-molecule freely orbiting magnetic tweezer (FOMT) method to investigate the handedness of DNA wrapping at a salt concentration of 50 mM KCl. Our experimental setup, analogous to that described in Vlijm *et al.* (*54*), is schematically shown in Fig. 6D. Observation of 11 independent assembly events shows that CAF-1 promoted right-handed DNA wrapping by an average of 0.9 turns, concomitant with the reduction of the DNA extension length by an average of ~19 nm (Fig. 6E and fig. S14A). By contrast, yeast NAP1 promotes left-handed DNA wrapping by an average of −0.9 turns, accompanied by the reduction of DNA extension of ~21 nm (fig. S14B). The latter result is consistent with a previous finding that NAP1-assembled tetrasomes are preferentially in a left-handed state, despite being able to spontaneously flip between the two handedness states (*54*). We cannot clearly distinguish whether DNA wraps one or two H3-H4 tetramers because the distribution of DNA extension change is quite broad, possibly indicating the existence of a mixture of tetrasomes and ditetrasomes.

Discussion

Our crystallographic and cryo-EM studies here reveal the long-anticipated structures of CAF-1 in the absence and presence of histones H3-H4. The structures uncover substantial structural plasticity and conformational dynamics underlying CAF-1's histone binding activity, and our biochemical, nucleosome assembly, and transcriptional analyses corroborate the structural findings that the ED domain of p150 is crucial for H3-H4 binding. The structure of the 1:1 complex of CAF-1 and H3-H4 shows that the homodimerization interface of H3 is exposed, but formation of an H3-H4 tetramer is prohibited because of steric clashes with p60. This finding agrees with and provides structural rationales to the recent results from biochemical analyses of the yeast complex (*35*, *52*).

Maintaining the integrity of parental and nascent H3-H4 tetramers during DNA replication has important implications in epigenetic inheritance (*4*). Our observation that the ED domain of p150 binds the large H3-H4 region that overlaps with the MCM2 binding site suggests a measure to segregate parental and newly synthesized H3-H4 (*40*). Assembly of H3-H4 tetramers from newly synthesized histones by CAF-1 is needed for proper chromatin inheritance. Interestingly, we found that the two H3-H4 heterodimers bound by two CAF-1 complexes are not in a tetrameric configuration, indicating that the process of H3-H4 tetramer assembly may be regulated. Although the structure of the 2:2 complex was only determined at relatively low resolution, the model was reliably built with the help of the higher-resolution 1:1 complex structure. In the 2:2 complex, the two H3-H4 heterodimers are positioned with the H3 dimerization interfaces next to each other, but their relative orientation and the distance between them deviate from that in an H3-H4 tetramer. It should be cautioned that our observation of the two H3-H4 dimers in the 2:2 complex does not exclude the possibility of a complex of CAF-1 with a H3-H4 tetramer that is not observed owing to low abundance or other conditions. For example, what we observed may represent a stable intermediate trapped under a specific condition, such as the DNA length used. In this arrangement, the spatial locations of the two p60s are the primary obstacle for aligning the two H3-H4 dimers for tetramer formation. Therefore, a remodeling of the interaction between p60 and H3-H4 is required for the assembly of an H3-H4 tetramer. It is possible that the binding of suitable-length DNA may facilitate efficient remodeling of the interaction between p60 and H3-H4 and the assembly of H3-H4 tetramers. Our SEC-MALS analysis showed that 30-bp DNA promoted the dimerization of two CAF-1–H3-H4 complexes, though the short DNA may have trapped the complexes at an intermediate state before the completion of the H3-H4 tetramer formation. Furthermore, a recent study also showed that DNA-loaded PCNA interacts with two CAF-1 complexes and promotes efficient nucleosome assembly (*55*).

Nucleosome assembly by CAF-1 occurs in two sequential steps: the formation of an H3-H4 tetrasome followed by the joining of H2A-H2B (*8*). Using the 147-bp Widom 601 DNA, we observed two different forms of ditetrasomes—a left-handed one and a CAF-1–bound right-handed one—from our in vitro assembly with H3-H4, DNA, and CAF-1. Considering that the left-handed ditetrasome was also obtained from the assembly of H3-H4 and DNA without CAF-1 (*53*), it most likely resulted from excessive H3-H4 added in our sample preparation. Handedness aside, the first critical question is whether ditetrasomes do occur during nucleosome assembly. Because histones H3-H4 are in a tetrameric form in the nucleosome, it has been taken for granted that the nucleosome assembly proceeds through an intermediate composed of a tetramer of H3-H4 wrapped by DNA. With a relatively short fragment of DNA (~80 bp), a well-defined complex of an H3-H4 tetramer with DNA can be identified in vitro (*35*, *56*). However, when presented with a longer DNA, the distinction between two independent H3-H4 tetrasomes or a ditetrasome is unclear. It also remains to be determined whether ditetrasomes, as a transient intermediate or as a relatively stable form of a non-nucleosomal histone-DNA complex, occur in a physiological setting. Nevertheless, one may imagine certain advantages of proceeding through a ditetrasome intermediate during nucleosome assembly. For example, a ditetrasome may serve as a more authentic "placeholder" for nucleosomes because it engages more than 100 to 120 bp of DNA instead of the ~45 to 55 bp for a tetrasome.

It is surprising that DNA wraps H3-H4 tetramers in a right-handed form in the CAF-1–bound ditetrasome. A previous study showed that CAF-1 is associated with the chromatin precursor (*8*), but the exact form of the precursor is unknown. Biophysical and biochemical studies demonstrated that H3-H4 tetrasomes exist in a dynamic equilibrium of left-handed and right-handed forms in vitro (*54*, *57*). However, very little is known about the right-handed DNA wrapping of histones in physiological settings, except in the context of centromeric chromatin (*58*). The structure of the CAF-1–bound right-handed ditetrasome and our single-molecule FOMT characterizations demonstrate that CAF-1 promotes right-handed DNA wrapping in vitro. Whether this also occurs in vivo and what the functional implication might be are not presently clear. We may speculate that a right-handed intermediate may balance the predominantly left-handed torsion of the tetrasome assembled from parental H3.1-H4, which are inherited mostly as an intact tetramer (*4*), to alleviate torsion stress and facilitate timely progression of the replication fork. It may also be helpful for replication of certain specialized chromosomal regions such as the centromere. Clearly, future efforts are needed to validate

and characterize the physiological roles of right-handed DNA wrapping by CAF-1.

Materials and methods

Plasmid construction

Recombinant CAF-1 subunits were produced in insect cells using baculoviruses prepared from modified vectors of pFastBac1 (Invitrogen). Plasmids carrying an N-terminal fusion of GST to p150 and a C-terminal fusion of 6×His tag to p60 and p48 were constructed by standard methods. A PreScission protease cleavage site was engineered between the GST tag and the p150 fragment. For large-scale production of full-length CAF-1 for structural and functional analyses, cDNA of full-length human p150 was subcloned into a modified pMlink vector containing N-terminal Protein A tag, and cDNAs for p60 and p48 were subcloned into pMlink vector without tags. A PreScission protease cleavage site was engineered between the Protein A tag and the p150 gene. Production of full-length human histones H3.1 and H4 in bacteria was performed using a bicistronic plasmid constructed with the pCDFDuet1 vector (Novagen). Truncation variants or point mutants of p150 and p60 and histones H3.1 or H4 were generated with the KOD-Plus-Mutagenesis kit (TOYOBO, SMK-101) following the manufacturer's protocols. All constructs were verified by DNA sequencing.

To convert the original AID system to AID2 (*44*), a point mutation was introduced into the OsTIR1 cDNA sequence in pMK232 (CMV-OsTIR1-PURO, Addgene, no. 72834). The OsTIR1(F74G) cDNA sequence was cloned into AAVS1-Neo-CAG-Flpe-ERT2 (Addgene, no. 68460) by Seamless Cloning (Beyotime) to replace the Flpe-ERT2 sequence. The spacer sequence targeting *AAVS1* was cloned into lentiCRISPRv2 (Addgene, no. 52961) following the instructions by the inventors. An AID-P2A-PuroR-tag was knocked-in immediately before the stop codon of *p150* or *p60* by CRISPR-Cas9–mediated homologous recombination. The construction of the repair template donor plasmid for the *p150*-AID knock-in was described previously (*59*), and the donor plasmid for the *p60*-AID knock-in was constructed accordingly. The spacer sequences targeting *p150* and *p60* were cloned into lentiCRISPRv2. For the p150 and p60 rescue experiments, full-length *p150* or *p60* cDNA was cloned into pCW-Cas9 (Addgene, no. 50661) by Seamless Cloning to replace the original spCas9 sequence, and blasticidin resistance was generated by replacing PuroR with blasticidin S deaminase sequence, which was subcloned from plasmid lenti dCAS-VP64_Blast (Addgene, no. 61425). p150 and p60 mutant constructs were generated by mutagenesis polymerase chain reaction (PCR).

Protein expression and purification

CAF-1 proteins were expressed in Sf21 cells using the Bac-to-Bac baculovirus expression system (Invitrogen). WT or various mutant CAF-1 complexes were produced by suspension cell culture coinfected with recombinant baculoviruses for each subunit following the manufacture's protocol. The full-length CAF-1 complex was also expressed using human embryonic kidney 293F (HEK293F) cells. A total of 500 ml of HEK293F cells was transfected with 0.5 mg of plasmids of three CAF-1 subunits (1:1:1 molar ratio) plus 2 mg of linear polyethylenimine. The cell culture was harvested 48 hours after transfection at 37°C.

The expressed CAF-1 middle domain complex (CAF1-MC) contains GST-tagged p150M (residues 442 to 714), p60ΔC-His, and p48-His, and the complex was purified first by resuspension of cell pellets in five volumes of the lysis buffer containing 20 mM Tris-HCl, pH 7.5, 500 mM NaCl, 5 mM imidazole, 1 mM phenylmethylsulphonyl fluoride (PMSF) and the protease inhibitor cocktail, before lysis by sonication. Cell debris was removed by centrifugation, and the cleared supernatant was incubated with Ni-NTA agarose resins for 30 min at 4°C, followed by washing the resins three times with the lysis buffer. The bound proteins were eluted with 500 mM imidazole in the lysis buffer. Immediately following the elution, 1 mM each of EDTA and dithiothreitol (DTT) were added to the Ni-NTA eluate, which was then incubated with glutathione (GSH) sepharose resins (GE Healthcare) preequilibrated in buffer A-500 (20 mM Tris-HCl, pH 7.5, 500 mM NaCl, 1 mM EDTA, and 1 mM DTT) for 2 hours. The loaded GSH resins were washed three times with buffer A-500, followed by elution of the bound proteins with the addition of 30 mM GSH. The GST-tag was cleaved by PreScission protease during dialysis to 100 mM NaCl overnight, and the sample was further purified on a Q column (GE Healthcare) with a linear 0.1 to 2 M NaCl gradient. Eluted fractions were analyzed by SDS–polyacrylamide gel electrophoresis (SDS-PAGE), and those highly enriched with CAF1-MC were pooled and concentrated and loaded onto a HiLoad 16/60 Superdex 200 column (GE Healthcare) in buffer A-500. Elution fractions containing highly purified CAF1-MC were pooled and concentrated to ~20 mg/ml by ultrafiltration and stored at −80°C before use.

The larger CAF-1 complex (CAF1-LC) containing the p150L (residues 442-853) fragment, which includes the WHD domain, was purified following a similar procedure, except that the ion-exchange column chromatography step was omitted. Full-length CAF-1 complex (CAF1-FC) supernatant was incubated with immunoglobulin G (IgG) beads for 2 hours, and unbound proteins were extensively washed away with buffer A-500. The fusion proteins were digested using PreScission protease overnight to remove tags, and the complex was eluted with buffer A-500. The eluted proteins were further purified using a Superose 6 increase 10/300 GL (GE Healthcare) size-exclusion column in buffer A-2000, which contains 2 M NaCl. C-terminal 6×His-tagged p48 and p60ΔC proteins were purified through successive Ni-NTA, respective ion-exchanges with Q and Heparin columns, and HiLoad 16/60 Superdex 200 sizing column steps.

WT and mutant human histone H3.1-H4 complexes were coexpressed using the pCDFDuet1-H3.1-H4 plasmid in *Escherichia coli* BL21 CodonPlus (DE3) RIL cells. Protein production was induced with 0.5 mM isopropyl-β-D-thiogalactopyranoside (IPTG) at 37°C for 4 hours. Histone complexes were first purified through a 5-ml SP column (GE Healthcare), followed by purification through a gel-filtration column at 2 M NaCl.

Crystallization

Initial small and thin-plate-shaped crystal clusters of the CAF1-MC complex were grown by sitting-drop vapor-diffusion at 20°C with 1.0 μl of protein solution (7 to 10 mg/ml) mixed with 1.0 μl of reservoir solution containing 8% (v/v) Tacsimate, pH 6.0, and 20% (w/v) PEG 3350. Multiple rounds of crystallization condition optimization yielded larger crystals in the reservoir solution containing 7% Tacsimate, pH 6.0, 22% (w/v) PEG 3350, 0.12 M lithium citrate, and 0.05 M sodium tartrate by hanging-drop vapor-diffusion at 4°C. Streak-seeding of previous optimized crystals into a protein-reservoir solution mixture at half of the protein concentration (3 to 5 mg/ml) preincubated overnight against reservoir solutions with 15 to 25% (w/v) PEG3350 and 0 to 0.2 M sodium tartrate was used to further improve the crystal quality. Single, thick crystals were obtained in a condition with 7% Tacsimate, pH 6.0, 18% (w/v) PEG 3350, 0.12 M lithium citrate, and 0.13 M sodium tartrate after 3 to 5 days. Cryogenic data collection was performed with cryoprotectants prepared by supplementing the crystallization mother liquor with 15% glycerol.

X-ray diffraction data collection, structure determination, and refinement

Diffraction data were collected at the Shanghai Synchrotron Radiation Facility (SSRF) beamline BL17U equipped with a Quantum 315r CCD detector (ADSC) using a wavelength of 0.97915 Å or beamline BL19U using Pilatus 6M detector at a wavelength of 0.97853 Å. Data were processed using the HKL2000 software package (*60*). Most of the crystals belong to the $C222_1$ space group, whereas a small percentage have a $P2_1$ or a C2 space group. The $C222_1$, $P2_1$, and C2 crystals diffracted to 3.5-, 3.6-, and 3.4-Å resolutions, respectively. The $C222_1$ structure was determined by molecular replacement (MR) with PHASER (*61*) using the crystal structures of human p48/RBBP4 [Protein Data Bank (PDB) ID 3GFC] and the human WD40-repeat protein Ciao1 (PDB ID 3FM0) as the search models.

Refinement of the MR solution gave rise to unambiguous extra densities attributable to p150 near p48 and p60. An initial model of CAF1-MC was manually built with COOT (*62*), and the model was improved by iterative cycles of refinement with PHENIX (*63*) and model adjustment. The refined model has R_{work} and R_{free} values of 0.225 and 0.257, respectively, and good stereochemical quality, with 96.8 and 3.2% of the residues in the favored and allowed regions of Ramachandran plot, respectively. The ordered structure contains residues 463 to 713 of p150, residues 1 to 394 of p60, and residues 7 to 411 of p48. Internal disordered segments of p150 spanning residues 607 to 657, which is highly enriched with acidic residues, and residues 100 to 111 of p60 are not included in the final model. The structures from the other two crystal forms were subsequently solved by molecular replacement using the $C222_1$ structure as a search model. Finally, the models were validated with MolProbity (*64*). Detailed statistics for data collection and structure refinement are shown in table S1.

DNA preparation

A 30-bp double-stranded DNA fragment (DNA_{30}) was prepared by annealing two high-performance liquid chromatography (HPLC)–grade chemically synthesized complementary DNA strands (forward: 5′-GTAATCCCCTTGGCGGTTAAAACGCGGGGG-3′; and reverse: 5′-CCCCCGCGTTTTAACCGCCAAGGGGATTAC-3′) in the annealing buffer (10 mM HEPES, pH 7.5, 100 mM NaCl, and 5 mM $MgCl_2$) at a concentration of 20 μM by ramping temperature down from 95° to 4°C at a rate of 0.1°C/s.

The 147-bp Widom 601 (W601) DNA was produced in *E. coli* following a published procedure (*65*, *66*). Briefly, tandem copies of the 147-bp W601 DNA fragment were inserted into a pEGFP-N1 vector (Addgene), and recombinant plasmids were produced in large bacterial culture. The 147-bp DNA was excised from the plasmid by EcoRV digestion, isolated by polyethylene glycol (PEG) precipitation, and further purified through ethanol precipitation. The sequence of the 147-bp W601 DNA is 5′-CTGGAGAATCCCGGTGCCGAGGCCGCTCAATTGGTCGTAGACAGCTCTAGCACCGCTTAAACGCACGTACGCGCTGTCCCCCGCGTTTTAACCGCCAAGGGGATTACTCCCTAGTCTCCAGGCACGTGTCACATATATACATCCTGT-3′.

Cryo-EM sample preparation

The CAF-1 complex with the p150L fragment encompassing the WHD domain (CAF1-LC) and histones H3-H4 were mixed at a 1:2.5 molar ratio in buffer A-500, followed by incubation on ice for 30 min, purification through a HiLoad 16/60 Superdex 200 column to remove excessive histones, and concentration of the purified complex to ~10 mg/ml. The preassembled CAF1-LC–H3-H4 complex and DNA_{30} were then mixed at a molar ratio of 1:1.5 in buffer A-500, and then dialyzed to buffer B (20 mM HEPES, pH 7.5, 50 mM NaCl, and 1 mM DTT) at 4°C for 36 hours. Preparation of the CAF1-FC–H3-H4-DNA30 complex is similar to that of the CAF1-LC complex, except that a Superose 6 increase 10/300 GL column (GE Healthcare) is used for the sizing column.

Preparation of the complexes of CAF-1, H3-H4, and W601 DNA (DNA_{147}) was performed by mixing WT or truncation variants of CAF-1, histones H3-H4, and DNA_{147} at a 1:2.5:1 molar ratio in buffer A-2000 and dialyzed to buffer B at 4°C in 36 hours.

The four samples above, with concentrations at ~1 mg/ml and a volume of 200 μl each, were subjected to glutaraldehyde cross-linking during GraFix (*67*). Specifically, the sample was subjected to a 12-ml linear 10 to 30% glycerol gradient in buffer B supplemented with 0 to 0.15% EM-grade glutaraldehyde (Sigma-Aldrich). After centrifugation at 4°C for 16 hours at 38,000 rpm in an SW40 rotor (Beckman Coulter), the sample was manually fractionated into 25 aliquots of 500 μl each from top to bottom using pipettes. The fractions were analyzed on an 8% denaturing polyacrylamide gel, and cross-linked fractions were further examined by negative-staining EM for particle homogeneity.

Suitable fractions were dialyzed to buffer A-50, which is identical to A-500 except that the NaCl concentration is changed to 50 mM, and cryo-EM specimens were prepared using the 300 mesh R2/1 CryoMatrix/Amorphous alloy film grid (Zhenjiang Lehua Technology Co, Ltd.) for the CAF1-LC–H3-H4–DNA_{30}, CAF1-FC–H3-H4–DNA_{30}, and CAF1-LC–H3-H4–DNA_{147} samples and the 300 mesh Quantifoil R2/1 gold grid for the CAF1-LC–H3-H4–DNA_{147} sample. The grids were treated with the Gatan Model 950 Advanced Plasma system using O_2/Ar for 60 s. A 3-μl aliquot of the sample at a concentration of ~0.4 mg/ml was applied to glow-discharged grids. After incubation for 10 s, excess sample was blotted with filter paper (diameter 55 mm, Whatman, GE Healthcare) for 3 to 4.5 s, and the grid was flash-frozen in liquid ethane using a FEI Vitrobot Mark VI device (ThermoFisher). All cryo-EM samples were prepared at 10°C and 100% relative humidity.

Cryo-EM data acquisition, image processing, and 3D reconstruction

Cryo-EM images were collected on an FEI Talos Arctica electron microscope equipped with a GIF Quantum energy filter and operating at 200 kV with a nominal magnification of 130,000. Images were recorded by a Gatan Bio-Quantum K2 Summit direct electron detector. The slit width for zero-loss peak was 20 eV. The camera was in a super-resolution mode with a 1-Å physical pixel size (0.5-Å super-resolution pixel size). The defocus range was set between −1.0 and −1.5 μm. Each image was exposed for 5 s, resulting in total electron exposure of ~50 $e^-/Å^2$ (exposure rate, 8.9 e^-/pixel/s) and 32 frames per movie stack.

A summary flow chart of data processing and structure determination procedures for the dataset obtained with the CAF1-LC–H3-H4 complex and DNA_{30} is shown in fig. S3. A total of 23,996 movies were collected and imported into cryoSPARC (*68*) and aligned using patch-based motion correction, and contrast transfer function (CTF) parameters were estimated in a patch manner using cryoSPARC and Gctf (*69*). Micrographs with poor Thon ring or heavy contamination were removed in the screening, and 19,607 micrographs were selected for further data processing. Particles of the complex were initially picked with cryoSPARC reference-free blob picker and used to generate 2D references for template-based auto-picking. Together with deep picking via topaz in cryoSPARC, a total of 19,430,612 particles were obtained. The dataset was cleaned up by several rounds of 2D classification, and 2,269,295 particles were kept for ab initio model generation (two classes) and subsequent heterogeneous refinement. These procedures resulted in two highly populated classes, monomeric and dimeric CAF1-LC–H3-H4 complexes. A new round of ab initio model generation (three classes for the monomer dataset and two for the dimer dataset), heterogeneous refinement, and nonuniform refinement on each class were performed to clean up the monomer and dimer datasets. For the monomer dataset, 927,863 particles in the best class were imported to Relion-3.0 (*70*) using the csparc2star.py script in pyem (*71*) and further sorted into four classes by 3D classification. The best class containing 400,801 particles was selected and further sorted into eight classes by no-alignment 3D classification. Three top classes were combined, and a total of 304,368 selected particles were subjected to cryoSPARC for nonuniform refinement, yielding a 3.9-Å map. Local refinement with a local mask improved the map resolution to 3.8 Å. For the dimer dataset, 53,708 particles were transferred to Relion-3.0 for 3D classification into four classes, and 24,904 particles in two classes were selected and subjected to nonuniform refinement, yielding a map with an overall resolution of 7.2 Å for the CAF1-LC–H3-H4 dimeric complex. Further refinement with C2 symmetry resulted in a 6.1-Å resolution map.

For the CAF1-FC–H3-H4–DNA_{30} dataset, as shown in fig. S4, a total of 15,893 movies were collected and imported into cryoSPARC and aligned using patch-based motion correction, and CTF parameters were estimated in a patch manner using cryoSPARC. Micrographs with poor Thon ring or heavy contamination were removed in the screening, and 14,532 micrographs were selected. The particles were picked with cryoSPARC reference-free blob picker initially and were used to generate 2D references for template-based autopicking. A total of 4,999,183

particles were autopicked using template and topaz. Several rounds of 2D classification were performed to clean up the dataset, and a total of 1,199,544 particles were kept for ab initio model generation and heterogeneous refinement using six classes. The best class with 688,535 particles was selected for nonuniform refinement, yielding a 3.6-Å map. Because the CAF1-FC–H3-H4 map is similar to that of CAF1-LC–H3-H4 (correlation value 0.976), particles in these two datasets were combined and overlaps were removed (within radius of 130 Å) to generate a combined set of 977,978 particles. Ab initio model generation and heterogeneous refinement using six classes resulted in two highly populated classes: monomers and dimers of CAF-1–H3-H4. The best two classes of the CAF-1–H3-H4 monomer containing 552,289 particles were selected and subjected to a further round of ab initio model generation and heterogeneous refinement using six classes. The best four classes containing 530,555 particles were selected for nonuniform refinement. Local refinement with a local mask improved the map resolution to 3.5 Å. In the final monomer particle set, 77 and 23% of the particles are from the FC and LC particle sets, respectively. The CAF1-FC dimer class containing 168,932 particles was combined with the CAF1-LC dimer dataset containing 53,708 particles, and the overlap removed within the 130-Å radius. A total of 197,148 particles were kept for ab initio model generation and heterogeneous refinement using six classes, and 50,508 particles in the best class were selected and subjected to nonuniform refinement, yielding a structure of the CAF-1–H3-H4 dimer at an overall resolution of 6.3 Å. Further refinement using C2 symmetry produced a 5.9-Å map of the dimeric CAF-1–H3-H4 complexes. Symmetry expanding to C1 generated 101,016 particles, and local refinements were performed by applying two local masks covering the two protomers, yielding two maps at 4.63-Å (protomer I) and 4.58-Å (protomer II) resolution. The composite map used for model building, refinement, and deposition was generated by aligning and merging the two locally refined maps using the Combine-Focused-Maps module in PHENIX. The half-maps of each local refined protomer were also merged and used for overall Fourier shell correlation (FSC) calculation in cryoSPARC, resulting in a 4.6-Å resolution composite map of the dimeric CAF-1–H3-H4 complex.

A summary flow chart of data processing and structure determination procedures for the peak 1 dataset of the CAF1-LC–H3-H4–DNA_{147} complex is shown in fig. S9. Data pre-processing steps were similar to those used with the 30-bp DNA oligomer. From this, 11,206 movies were collected and 10,316 micrographs were selected for particle picking after patch motion correction and patch CTF estimation by cryoSPARC. Because of the relatively small size of the protein complex, it is difficult to distinguish particles of the intact protein complex or partial complexes, and a large set of particles, 10,865,604 in total, were initially picked and extracted. The dataset was cleaned up through several rounds of 2D classification, and a total of 1,088,125 particles were kept for ab initio model generation and heterogeneous refinement. Two highly populated classes were obtained: one is a left-handed ditetrasome, and the other is the CAF-1–bound right-handed ditetrasome. These two subsets of data were imported to Relion-3.0 for further 3D classification.

For the left-handed ditetrasome subset of data, 459,354 particles were sorted into four classes by 3D classification. Three top classes were combined, and a total of 335,568 particles were selected for nonuniform refinement, which yielded a map at an overall resolution of 3.5 Å. For the CAF-1–bound right-handed ditetrasome subset of data, 628,771 particles were sorted into six classes by 3D classification. Two top classes containing 178,734 particles were combined and subjected to nonuniform refinement, yielding a map at an overall resolution of 3.8 Å.

Figure S10 summarizes data processing and structure determination procedures for the dataset collected from the peak 2 sample. From this, 2213 movies were collected, and 2007 micrographs were kept for particle picking. A large set of 1,568,202 particles were initially picked and extracted. A total of 243,993 particles were kept after multiple rounds of 2D classification and used for ab initio model generation and heterogeneous refinement. Three classes of particles were obtained, and the 126,187 particles in the best class were imported to Relion-3.0 for 3D classification and resulted in six classes. One class containing 8190 particles showed two extra chunks of density on one side of the ditetrasome surface, with a low resolution of ~18 Å. Two other classes both have one extra piece of density on each side of the ditetrasome surface, and the better class containing 76,910 particles was selected and subjected to nonuniform refinement, yielding a map of two CAF-1–bound ditetrasomes at an overall resolution of 5.6Å and 6.3 Å with and without C2 symmetry imposed, respectively.

A flow chart summarizing data processing procedures for the CAF1-FC–H3-H4–DNA_{147} complex is shown in fig. S11. From this, 8801 movies were collected, and 8002 micrographs were kept for particle picking. A large set of 3,199,346 particles were initially picked and extracted. A total of 947,480 particles were kept after multiple rounds of 2D classification and used for ab initio model generation and heterogeneous refinement with six classes. The 104,027 particles in the best class were further used for the next round of ab initio model generation and heterogeneous refinement by six classes. These procedures resulted in two highly populated classes: single and double CAF1-FC–ditetrasome complexes. For the single CAF1-FC–ditetrasome, two classes containing 67,319 particles were subjected to nonuniform refinement, yielding a map of single CAF-1–bound ditetrasome at an overall resolution of 3.8 Å. The 13,993 particles of the double CAF1-FC–ditetrasome class were used for nonuniform refinement, yielding a map of the double CAF-1–bound ditetrasome at an overall resolution of 6.6 Å.

All reported resolutions were estimated based on the gold-standard FSC criterion of 0.143. Local resolutions were estimated by cryoSPARC.

Model building and refinement of the cryo-EM structures

Crystal structures of CAF1-MC from this study and the histone H3.1-H4 model from the sNASP–ASF1–H3.1-H4 complex (PDB ID 7V6Q) were used as initial models for building the model of the CAF1-LC–H3-H4 complex by docking each component into the cryo-EM map in UCSF Chimera (*72*). The resulting 1:1 model of the CAF1-LC–H3-H4 complex was used as the initial model for fitting the structure of the 2:2 complex. For the left-handed ditetrasome structure, the atomic models of the Widom 601 DNA (PDB ID 3LZ0) and the histone H3.1-H4 heterodimer were fitted into the cryo-EM map. For the CAF-1–bound ditetrasome, the crystal structure CAF1-MC from this study and histones H3.1-H4 were fitted into the map, and the right-handed Widom 601 DNA model was built manually. An additional CAF1-MC complex was introduced into the model of the right-handed ditetrasome bound by two CAF-1 complexes.

The final models were obtained after iterative cycles of real-space refinement using PHENIX with secondary structure and Ramachandran and rotamer restraints, assisted with manual editing, adjustment and model rebuilding in COOT. Three macro cycles of final refinement were performed with the minimization_global and local_grid_search strategies, and the qualities of the refined models were checked with the MolProbity parameters calculated using PHENIX. For the four low-resolution structures, the 4.6- and the 6.1-Å dimeric CAF-1–H3-H4 complexes and the 5.6- and the 6.6-Å double CAF-1–bound ditetrasome complexes, only main chains were kept in these models. These models were real-space refined using PHENIX as rigid bodies with secondary structure and Ramachandran restraints. Also, in the two 3.8-Å single CAF-1–ditetrasome complexes, only the main-chain model for p60 and p150 were kept because of the limit of local resolutions. The final refinement statistics are shown in tables S2 and S3. Structural figures were prepared using Chimera, ChimeraX (*73*), and PyMOL (Schrödinger LLC, New York, NY).

Topological assay for nucleosome assembly

DNA supercoiling assays were performed as described previously (*74*) with the following

modifications. Relaxed DNA template was prepared by combining ϕX174 RF I DNA (NEB, N3021S) at 100 ng per sample, with excess DNA Topoisomerase I (TaKaRa, 2240A) in the assembly buffer [10 mM Tris-HCl, pH 7.5, 125 mM NaCl, 2 mM $MgCl_2$, 0.5 mM DTT, and 0.1 mg/ml bovine serum albumin (BSA)] at 37°C for 1 hour and kept at room temperature ready for use. Histones H3.1-H4 (100 ng) were preincubated with 100 and 200 ng each of yNAP1, WT CAF-1, or its mutants in the assembly buffer at 37°C for 30 min. To initiate the assembly reaction, 100 ng of relaxed plasmid DNA was added to the chaperone-histone mixture and incubated for 15 min at 37°C. The 25-μl standard chromatin assembly reaction was stopped by the addition of EDTA to 20 mM and SDS to 0.5% and 0.5 mg/ml proteinase K, followed by incubation at 55°C for 15 min before phenol-chloroform extraction. The DNA samples were analyzed on a 1% agarose gel in 1× TAE buffer (40 mM Tris-acetate, 1 mM EDTA) at 90 V for 4 hours with ethidium bromide staining.

GST-pulldown

To analyze the interaction of p150 with p48 and p60ΔC proteins, 3 μg of GST or GST-fused p150M or different p150 fragments were first immobilized on 5 μl of glutathione Sepharose 4B resins (GE Healthcare), followed by incubation with 5 μg each of p48 and p60ΔC in 500 μl of binding buffer (20 mM Tris-HCl, pH 7.5, 500 mM NaCl, 5 mM 2-mercaptoethanol, and 1 mM EDTA) under continuous rotation at 4°C for 2 hours. The resins were then washed five times with 1 ml of binding buffer. Bound proteins were eluted in the SDS sample buffer, separated on a 15% SDS-PAGE gel, and stained with Coomassie blue.

To test the interactions between CAF-1 and histones H3-H4, 3 μg of GST or GST-fused WT or mutant CAF-1 complexes were immobilized on 5 μl of glutathione Sepharose 4B resin (GE Healthcare), which was then mixed with 5 μg of H3.1-H4 or its mutants in 500 μl of binding buffer (20 mM Tris-HCl, pH 7.5, 500 mM NaCl, 5 mM 2-mercaptoethanol, 0.01% NP-40, and 1 mM EDTA), and the mixture was rotated overnight at 4°C. The resins were then washed five times with 1 ml of washing buffers modified from the binding buffer with 0.1% NP-40 and 500 mM or 1 M NaCl. Bound proteins were eluted in the SDS sample buffer and analyzed by Coomassie-stained SDS-PAGE.

Cell lines

HEK 293FT cells (Thermo Fisher, R70007) were maintained in Dulbecco's modified Eagle medium (DMEM) (Gibco) supplemented with 10% (v/v) fetal bovine serum (FBS) (Biological Industries) and penicillin-streptomycin (Sangon Biotech). HAP1 cells (Horizon) were cultured in Iscove's modified Dulbecco's medium (IMDM) (Gibco) supplemented with 10% FBS and penicillin-streptomycin. To integrate the OsTIR1 expression cassette into the *AAVS1* locus, cells cotransfected with AAVS1-Neo-CAG-OsTIR1(F74G) and lentiCRISPR-sgAAVS1 by jetOPTIMUS (Polyplus) were selected by 1 mg/ml G418 for 7 days and then seeded into 96-well plates by fluorescence-activated cell sorting (FACS). Clones from single cells were expanded and analyzed by Western blotting and reverse transcription quantitative PCR (RT-qPCR). AID-tagged *p150* or *p60* fusion clones were generated as previously described (*75*). Transfected cells were selected by 0.2 μg/ml puromycin for 2 days. Cell clones from single cells were expanded and analyzed by genotyping PCR and Western blotting. Only clones with biallelic insertions were kept for future experiments. Constructs expressing *p150* or *p60* were introduced into p150-AID or p60-AID cells by lentiviral transduction, respectively; stable cells were selected by 5 μg/ml blasticidin; and expressions of exogenously introduced genes were examined by Western blotting.

RNA-seq and data analysis

Cells grown in a six-well plate were extracted with 1 ml of TRIzol reagent (ThermoFisher, 15596026) according to the manufacturer's instructions. Total RNA was prepared with two biological replicates for each genotype or treatment. Poly-A–tailed mRNA molecules were captured using attached oligo-dT and subjected to library construction according to the manufacturer's recommendations.

mRNA-seq libraries were sequenced on the NovaSeq 6000 platform by Annoroad Gene Technology Co., Ltd., using STAR aligner (v2.7.9a), and mRNA-seq read pairs were mapped to the human genome sequences (hg38) with Gencode annotations. Differential gene expression analysis was performed using DESeq2 software, and differentially expressed genes (DEGs) were identified using the cutoff adjusted p value ≤0.05 and fold change ≥2.

Cell proliferation assays

To analyze the effects of p150 and p60 mutants on cellular proliferation, cells were seeded into 96-well plates at a density of 1000 cells per well. Doxycycline was added to the medium at a final concentration of 2 μg/ml while seeding. After 24 hours, the medium was changed, and 5-Ph-IAA was added to cells at a final concentration of 1 μM to induce the degradation of endogenous p150 or p60 in the presence of doxycycline to maintain the expression of exogenous p150 or p60. Medium was changed every day to supply fresh doxycycline and 5-Ph-IAA. At 6 days after 5-Ph-IAA treatment, cell viability was measured by CellTiter-Glo reagent (Promega). Survival rate was quantified as the ratio of the number of survived cells to the cell number in the control treatment.

ReIN-Map, MNase-seq, and data analysis

The ReIN-Map protocol was adapted and modified from those used in (*76*–*78*). To analyze the effects of CHAF1A and CHAF1B mutants on nascent nucleosome landscape after DNA replication, cells were seeded into 150-mm dishes at a density of 1×10^7 per dish. Twelve hours after seeding, doxycycline was added to the medium at a final concentration of 2 μg/ml. After 24 hours, the medium was changed, and 5-Ph-IAA was added to a final concentration of 1 μM in the presence of doxycycline to maintain the expression of exogenous CHAF1A or CHAF1B. After 6 hours, the cells were pulse-labeled with 20 μM EdU (SANTA CRUZ, sc-284628) for 10 min. After the EdU pulsing, cells were immediately fixed in 1% formaldehyde for 10 min. Then, glycine was added to a final concentration of 0.125 M, and the reaction was incubated for 5 min at room temperature. Cells were then collected by scraping with a cell lifter and transferred into 15-ml tubes. After centrifugation, the cell pellets were resuspended in cold PBS and transferred into a 1.5-ml Eppendorf tube.

Fixed cells were resuspended in Buffer I (50 mM HEPES, pH 7.5, 1 mM EDTA, 140 mM NaCl, 10% glycerol, 0.5% NP-40, and 0.25% Trition X-100) and incubated at 4°C for 10 min. After centrifugation, the cell pellets were resuspended in Buffer II (10 mM Tris-HCl, pH 7.5, 1 mM EDTA, 0.5 mM EGTA, and 200 mM NaCl) and incubated at room temperature for 10 min. After centrifugation, cell nuclei were resuspended with Buffer III (10 mM Tris-HCl, pH 7.5, 60 mM KCl, 15 mM NaCl, and 3 mM $MgCl_2$), and briefly sonicated for 30 s with a bioruptor (Diagenode). The samples were split into two 1.5-ml Eppendorf tubes for MNase digestion and chromatin sonication separately. For MNase digestion, 300-μl samples were incubated at 37°C for 3 min with 10 μl of MNase (lab stock), and digested in the presence of 2 mM $CaCl_2$ for 10 min. The digestion was stopped with a final concentration of 10 mM EDTA, and SDS was added at a final concentration of 1%. For chromatin sonication, SDS was added at a final concentration of 1%, and 300-μl samples were transferred into a sonication tube. Then, 1 ml of NLB buffer (50 mM Tris, pH 7.5, 10 mM EDTA, and 1% SDS) and sonication beads were added to each sample and sheared to 300- to 500-bp-sized fragments with a bioruptor. Both MNase-treated and sonication-fragmented chromatin were de–cross-linked at 65°C overnight with 50 μg of proteinase K. After incubation, phenol-chloroform-isoamyl alcohol (PCI) extraction was performed, and RNase-A was added to the aqueous phase, which was then incubated at 37°C for 15 min. PCI extraction was performed again, and the DNA in the aqueous phase was precipitated by adding 2.5× volume of ethanol, 200 mM NaCl, and glycogen and then incubated at −20°C overnight. After centrifugation, the DNA pellet was washed with 70% ethanol and resuspended in 10 mM Tris-HCl, pH 7.5. The extracted DNA

fragments were subjected to end repair, A-tailing, and adaptor ligation using the NEB Ultra II DNA Library Prep kit (E7645) following the manufacturer's instructions. After reserving an aliquot for input (MNase-seq), EdU-pulsed DNA (MINCE-seq and Sonication-seq) was biotinylated under the following conditions: 15 mM Tris-HCl, pH 7.5, 0.5 mM biotin-TEG-azide (Jena Bioscience), 0.1 mM $CuSO_4$ (Sigma-Aldrich), 0.5 mM THPTA (Sigma-Aldrich), and 10 mM sodium ascorbate (Sigma-Aldrich) for 45 min at room temperature. DNA fragments were purified using AMPure XP beads (1.1×) and resuspended in 10 mM Tris-HCl, pH 7.5. Next, biotinylated DNA was pulled down using Myone C1 beads (Invitrogen). Six μl of beads per sample were balanced by three washes with 1× BW&T buffer (5 mM Tris-HCl pH 7.5, 0.5 mM EDTA, 1 M NaCl, and 0.05% Tween-20) and resuspended in 2× BW&T buffer. Streptavidin beads were then mixed 1:1 with biotinylated DNA and rotated for 45 min at room temperature. Then, the beads were washed four times with 1× BW&T, twice with 1× TE with 0.05% Tween-20, and once with 10 mM Tris-HCl, pH 7.5 and resuspended in 10 μl of ddH_2O. The beads were incubated at 98°C for 10 min, and 16 cycles of PCR amplification were performed together with input DNA libraries. Post-PCR clean-up was performed by adding 1× volume of AMPure XP beads. Two independent biological repeats were performed for each sample, and similar results were obtained for all analyses.

Pair-end sequencing reads were processed, and the samples were mapped to *Homo sapiens* genome (hg38) by Bowtie2 (*79*) using default parameters. Uniquely mapped reads were collected for further analysis. Nucleosome positions were called using the statistic pipeline DANPOS (*80*) with the fragments of size 120 to 180 bp. Mononucleosome coverage was smoothed with a 10-bp flat window. ReIN score is defined as the normalized read density of MINCE-seq divided by the density of Sonication-seq samples. MINCE-seq signals are influenced by both DNA synthesis and nucleosome assembly. Thus, by normalizing MINCE-seq with Sonication-seq, the nucleosome occupancy score reflects nucleosome occupancy independent of DNA synthesis and EdU incorporation. All profiles relative to a feature of TSSs or CTCF-binding sites were plotted as the ReIN score or nucleosome occupancy score spanning a given position relative to the feature over a ±1000-bp window around the feature midpoint. The CTCF chromatin immunoprecipitation sequencing (ChIP-seq) data are from the Gene Expression Omnibus (GEO) (GSM4640493) (*81*)

Molecular weight determination by SEC-MALS

Molar masses of protein or protein-DNA complexes in solution were determined by SEC-MALS using the DAWN HELEOSTM II 18-angle static light-scattering system (Wyatt Technology) connected to an Agilent HPLC that was hooked up with a Superose 6 increase 10/300 GL column (GE Healthcare). The system was first equilibrated in a buffer containing 20 mM Tris-HCl, pH 7.5, 125 mM NaCl, and 2% glycerol for 12 hours. The equilibrated system was then calibrated with BSA at a concentration of 1 mg/ml. Fifty microliters of purified CAF-1 or CAF-1–H3-H4 complexes in the presence and absence of DNA_{30} at a concentration of ~1 mg/ml each were subjected to the system at a flow rate of 0.5 ml/min at room temperature. Data were analyzed by the ASTRA software.

Single-molecule FOMT analysis

A multiple digoxigenin-labeled and multiple biotin-labeled 159-bp DNA segment was amplified by PCR and ligated with handle DNA (543 bp) via the *Sty*I restriction site. Purified ligation products were further ligated with the 147-bp Widom 601 DNA via the *Bsa*I restriction site, and the full-ligation product was purified by gel extraction.

Upon treating the coverslips with piranha solution, polystyrene beads (2 μm, QDSpher) were loaded and incubated at 150°C for 8 min on a hot plate, and the beads served as the reference to control thermal drifts during the experiment. Then, the coverslips were coated with Sigmacote (Sigma-Aldrich, SL2) for 5 min and the flow-cells were incubated with 100 μl of antidigoxigenin (0.1 mg/ml, Roche) in PBS for 4 hours at 4°C. This was followed by passivation of the flow-cell with 100 μl of passivation buffer [10 mg/ml BSA, 1 mM EDTA, 10 mM phosphate buffer, pH 8.0, 10 mg/ml Pluronic F127 surfactant (Sigma-Aldrich, P2443), and 3 mM NaN_3] at 4°C overnight.

FOMTs (*54*, *82*), as schematically shown in Fig. 6D, were built based on homemade single-molecule magnetic tweezers described before (*83*). Force calibration was carried out by tethering an 8-kb DNA between the superparamagnetic bead (Dynabeads MyOne Streptavidin T1, Invitrogen, 65601) and the coverslip, following published procedures (*82*, *84*), with the recording of freely rotating magnetic beads positioned under a force ranging from 0.45 to 1.8 pN.

After the flow-cell was rinsed with 1 ml of FOMT buffer (50 mM KCl, 25 mM HEPES, pH 7.5, 0.1 mM EDTA, 0.025% PEG, 0.025% PVOH, and 1 mg/ml BSA), 100 μl of purified full-ligated DNA (3 pg/μl) in FOMT buffer was loaded and incubated at room temperature for 20 min before the flow-cell was further washed with 200 μl of FOMT buffer. Superparamagnetic bead (T1) in FOMT buffer was added to the flow-cell and incubated for 20 min at room temperature. Then, 200 μl of FOMT buffer was flowed into the flow-cell to flush away the unbounded beads.

Yeast Nap1-H3H4 (200 nM) and human CAF1-FC–H3-H4 (50 nM) complexes were kept in the preincubation buffer (50 mM KCl, 25 mM HEPES, pH 7.5, 0.1 mM EDTA, 0.25% PEG, 0.25% PVOH, and 1 mg/ml BSA) at 4°C for 30 min. Then the complexes were loaded into the flow-cell with 100 times dilution. The bead image was captured by a JAI Giga-Ethernet CCD camera at 500 Hz for 4 hours through the inverted microscope objective lens (UPLXAPO60XO, NA 1.42, Olympus). The real-time position (x, y, z) of the beads at 0.7 pN was recorded with the comparison between the diffraction pattern of the beads and calibration images at various distances from the focal point of the objective.

The circle center (x_c, y_c) and radius R_{circle} of the beads' trajectories are fitted using the least squares method, that is, minimizing the sum of the squares of the residuals:

$$\min_{x_c, y_c, R_{circle}} = \sum_i \left[(x_i - x_c)^2 + (y_i - y_c)^2 - R_{circle}{}^2\right] \quad (1)$$

To calculate parameters x_c, y_c, and R_{circle} from the bead's position (x_i, y_i), we defined matrix A and vector $\vec{b}, \vec{x}$:

$$A = \begin{pmatrix} x_1 & y_1 & 1 \\ x_1 & y_2 & 1 \\ \cdots & \cdots & \cdots \\ x_n & y_n & 1 \end{pmatrix},$$

$$\vec{b} = \begin{pmatrix} x_1{}^2 + y_1{}^2 \\ x_2{}^2 + y_2{}^2 \\ \cdots \\ x_n{}^2 + y_n{}^2 \end{pmatrix}, \text{and}$$

$$\vec{x} = \begin{pmatrix} 2x_c \\ 2y_c \\ r^2 - x_c{}^2 - y_c{}^2 \end{pmatrix} \quad (2)$$

The overdetermined system is

$$A\vec{x} = \vec{b} \quad (3)$$

The solution of which is given by

$$\vec{x} = \left(A^{\mathrm{T}}A\right)^{-1}A^{\mathrm{T}}\vec{b} \quad (4)$$

After finding x_c, y_c, and R_{circle}, we set the pole of the polar coordinates system as (x_c, y_c). By denoting the bead's position vector (x_i, y_i) as $\vec{r}_i$, the rotation from position $\vec{r}_i$ to $\vec{r}_{i+1}$ is given by

$$\Delta\theta_{i,i+1} = \arccos\frac{\left\langle \vec{r}_i, \vec{r}_{i+1} \right\rangle}{\left|\vec{r}_i\right|\left|\vec{r}_{i+1}\right|} \quad (5)$$

The angles of rotation from the initial position are given by

$$\theta_k = \begin{cases} \sum_{i=1}^{k-1} \Delta\theta_{i,i+1}, & \Delta\theta_{i,i+1} \\ 0, & k = 1 \end{cases} \quad (6)$$

The polar coordinates (r_i, ϕ_i) are given by

$$r_i = \sqrt{(x_i - x_c)^2 + (y_i - y_c)^2} \quad (7)$$

$$\phi_i = \arctan\left(\frac{y_i - y_c}{x_i - x_c}\right) \quad (8)$$

The code used for analyzing and plotting FOMT data is archived at Zenodo (*85*).

EMSA

In each EMSA reaction (total volume of 20 µl), a mixture of 0.5 µM 147-bp DNA with 0.5, 1, 2.5, 5, or 10 µM of p60ΔC or p48, in a buffer containing 20 mM Tris-HCl, pH 7.5, 50 mM NaCl, and 1 mM DTT, was first incubated on ice for 30 min, followed by electrophoresis on a 6% native polyacrylamide gel at 120 V for 50 min with 0.5× TBE buffer in an ice-water bath. The gel was stained by SYBR Gold (Thermo Fisher) and visualized with the Gel Doc EZ imaging System (Bio-Rad).

Quantification and statistical analysis

The orientation distribution of the particles used in the final reconstructions and the local resolution maps were calculated using cryoSPARC (*68*). The quantification and statistical analyses for model refinement and validation were generated using MolProbity (*64*). Data for cell proliferation assays are from three biological replicates, and a two-way analysis of variance (ANOVA) with Sídák test was used to calculate p values. Data for RNA-seq are presented with two biological replicates.

REFERENCES AND NOTES

1. K. Luger, A. W. Mäder, R. K. Richmond, D. F. Sargent, T. J. Richmond, Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature* **389**, 251–260 (1997). doi: 10.1038/38444; pmid: 9305837
2. A. Serra-Cardona, Z. Zhang, Replication-coupled nucleosome assembly in the passage of epigenetic information and cell identity. *Trends Biochem. Sci.* **43**, 136–148 (2018). doi: 10.1016/j.tibs.2017.12.003; pmid: 29292063
3. A. Groth, W. Rocha, A. Verreault, G. Almouzni, Chromatin challenges during DNA replication and repair. *Cell* **128**, 721–733 (2007). doi: 10.1016/j.cell.2007.01.030; pmid: 17320509
4. M. Xu *et al*., Partitioning of histone H3-H4 tetramers during DNA replication-dependent chromatin assembly. *Science* **328**, 94–98 (2010). doi: 10.1126/science.1178994; pmid: 20360108
5. C. Yu *et al*., A mechanism for preventing asymmetric histone segregation onto replicating DNA strands. *Science* **361**, 1386–1389 (2018). doi: 10.1126/science.aat8849; pmid: 30115745
6. N. Petryk *et al*., MCM2 promotes symmetric inheritance of modified histones during DNA replication. *Science* **361**, 1389–1392 (2018). doi: 10.1126/science.aau0294; pmid: 30115746
7. S. Smith, B. Stillman, Purification and characterization of CAF-I, a human cell factor required for chromatin assembly during DNA replication in vitro. *Cell* **58**, 15–25 (1989). doi: 10.1016/0092-8674(89)90398-X; pmid: 2546672
8. S. Smith, B. Stillman, Stepwise assembly of chromatin during DNA replication in vitro. *EMBO J.* **10**, 971–980 (1991). doi: 10.1002/j.1460-2075.1991.tb08031.x; pmid: 1849080
9. P. D. Kaufman, R. Kobayashi, N. Kessler, B. Stillman, The p150 and p60 subunits of chromatin assembly factor I: A molecular link between newly synthesized histones and DNA replication. *Cell* **81**, 1105–1114 (1995). doi: 10.1016/S0092-8674(05)80015-7; pmid: 7600578
10. Z. Zhang, K. Shibahara, B. Stillman, PCNA connects DNA replication to epigenetic inheritance in yeast. *Nature* **408**, 221–225 (2000). doi: 10.1038/35041601; pmid: 11089978
11. K. Shibahara, B. Stillman, Replication-dependent marking of DNA by PCNA facilitates CAF-1-coupled inheritance of chromatin. *Cell* **96**, 575–585 (1999). doi: 10.1016/S0092-8674(00)80661-3; pmid: 10052459
12. J. K. Tyler *et al*., The RCAF complex mediates chromatin assembly during DNA replication and repair. *Nature* **402**, 555–560 (1999). doi: 10.1038/990147; pmid: 10591219
13. J. A. Sharp, E. T. Fouts, D. C. Krawitz, P. D. Kaufman, Yeast histone deposition protein Asf1p requires Hir proteins and PCNA for heterochromatic silencing. *Curr. Biol.* **11**, 463–473 (2001). doi: 10.1016/S0960-9822(01)00140-3; pmid: 11412995
14. M. Hoek, B. Stillman, Chromatin assembly factor 1 is essential and couples chromatin assembly to DNA replication in vivo. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 12183–12188 (2003). doi: 10.1073/pnas.1635158100; pmid: 14519857
15. S. Enomoto, P. D. McCune-Zierath, M. Gerami-Nejad, M. A. Sanders, J. Berman, RLF2, a subunit of yeast chromatin assembly factor-I, is required for telomeric chromatin function in vivo. *Genes Dev.* **11**, 358–370 (1997). doi: 10.1101/gad.11.3.358; pmid: 9030688
16. P. D. Kaufman, R. Kobayashi, B. Stillman, Ultraviolet radiation sensitivity and reduction of telomeric silencing in *Saccharomyces cerevisiae* cells lacking chromatin assembly factor-I. *Genes Dev.* **11**, 345–357 (1997). doi: 10.1101/gad.11.3.345; pmid: 9030687
17. Q. Li *et al*., Acetylation of histone H3 lysine 56 regulates replication-coupled nucleosome assembly. *Cell* **134**, 244–255 (2008). doi: 10.1016/j.cell.2008.06.018; pmid: 18662540
18. H. Tagami, D. Ray-Gallet, G. Almouzni, Y. Nakatani, Histone H3.1 and H3.3 complexes mediate nucleosome assembly pathways dependent or independent of DNA synthesis. *Cell* **116**, 51–61 (2004). doi: 10.1016/S0092-8674(03)01064-X; pmid: 14718166
19. K. Ahmad, S. Henikoff, The histone variant H3.3 marks active chromatin by replication-independent nucleosome assembly. *Mol. Cell* **9**, 1191–1200 (2002). doi: 10.1016/S1097-2765(02)00542-7; pmid: 12086617
20. E. Martini, D. M. Roche, K. Marheineke, A. Verreault, G. Almouzni, Recruitment of phosphorylated chromatin assembly factor 1 to chromatin after UV irradiation of human cells. *J. Cell Biol.* **143**, 563–575 (1998). doi: 10.1083/jcb.143.3.563; pmid: 9813080
21. P. Ridgway, G. Almouzni, CAF-1 and the inheritance of chromatin states: At the crossroads of DNA replication and repair. *J. Cell Sci.* **113**, 2647–2658 (2000). doi: 10.1242/jcs.113.15.2647; pmid: 10893180
22. S. Cheloufi, K. Hochedlinger, Emerging roles of the histone chaperone CAF-1 in cellular plasticity. *Curr. Opin. Genet. Dev.* **46**, 83–94 (2017). doi: 10.1016/j.gde.2017.06.004; pmid: 28692904
23. A. Verreault, P. D. Kaufman, R. Kobayashi, B. Stillman, Nucleosome assembly by a complex of CAF-1 and acetylated histones H3/H4. *Cell* **87**, 95–104 (1996). doi: 10.1016/S0092-8674(00)81326-4; pmid: 8858152
24. P. V. Sauer *et al*., Mechanistic insights into histone deposition and nucleosome assembly by the chromatin assembly factor-1. *Nucleic Acids Res.* **46**, 9907–9917 (2018). doi: 10.1093/nar/gky823; pmid: 30239791
25. T. Rolef Ben-Shahar *et al*., Two fundamentally distinct PCNA interaction peptides contribute to chromatin assembly factor 1 function. *Mol. Cell. Biol.* **29**, 6353–6365 (2009). doi: 10.1128/MCB.01051-09; pmid: 19822659
26. A. Volk, J. D. Crispino, The role of the chromatin assembly complex (CAF-1) and its p60 subunit (CHAF1b) in homeostasis and disease. *Biochim. Biophys. Acta* **1849**, 979–986 (2015). doi: 10.1016/j.bbagrm.2015.05.009; pmid: 26066981
27. S. E. Polo *et al*., Chromatin assembly factor-1, a marker of clinical value to distinguish quiescent from proliferating cells. *Cancer Res.* **64**, 2371–2381 (2004). doi: 10.1158/0008-5472.CAN-03-2893; pmid: 15059888
28. K. S. Shim *et al*., Reduction of chromatin assembly factor 1 p60 and C21orf2 protein, encoded on chromosome 21, in Down syndrome brain. *J. Neural Transm. Suppl.* **67**, 117–128 (2003). doi: 10.1007/978-3-7091-6721-2_10; pmid: 15068244
29. Y. Tang *et al*., Structure of a human ASF1a-HIRA complex and insights into specificity of histone chaperone complex assembly. *Nat. Struct. Mol. Biol.* **13**, 921–929 (2006). doi: 10.1038/nsmb1147; pmid: 16980972
30. F. Mattiroli, Y. Gu, J. L. Balsbaugh, N. G. Ahn, K. Luger, The Cac2 subunit is essential for productive histone binding and nucleosome assembly in CAF-1. *Sci. Rep.* **7**, 46274 (2017). doi: 10.1038/srep46274; pmid: 28418026
31. J. K. Tyler, M. Bulger, R. T. Kamakaka, R. Kobayashi, J. T. Kadonaga, The p55 subunit of *Drosophila* chromatin assembly factor 1 is homologous to a histone deacetylase-associated protein. *Mol. Cell. Biol.* **16**, 6149–6159 (1996). doi: 10.1128/MCB.16.11.6149; pmid: 8887645
32. A. Verreault, P. D. Kaufman, R. Kobayashi, B. Stillman, Nucleosomal DNA regulates the core-histone-binding subunit of the human Hat1 acetyltransferase. *Curr. Biol.* **8**, 96–108 (1998). doi: 10.1016/S0960-9822(98)70040-5; pmid: 9427644
33. Y. Yue, W. S. Yang, L. Zhang, C. P. Liu, R. M. Xu, Topography of histone H3-H4 interaction with the Hat1-Hat2 acetyltransferase complex. *Genes Dev.* **36**, 408–413 (2022). doi: 10.1101/gad.349099.121; pmid: 35393344
34. D. Kim *et al*., Molecular architecture of yeast Chromatin Assembly Factor 1. *Sci. Rep.* **6**, 26702 (2016). doi: 10.1038/srep26702; pmid: 27221973
35. P. V. Sauer *et al*., Insights into the molecular architecture and histone H3-H4 deposition mechanism of yeast Chromatin assembly factor 1. *eLife* **6**, e23474 (2017). doi: 10.7554/eLife.23474; pmid: 28315525
36. K. Zhang *et al*., A DNA binding winged helix domain in CAF-1 functions with PCNA to stabilize CAF-1 at replication forks. *Nucleic Acids Res.* **44**, 5083–5094 (2016). doi: 10.1093/nar/gkw106; pmid: 26908650
37. H. Wang, M. Wang, N. Yang, R. M. Xu, Structure of the quaternary complex of histone H3-H4 heterodimer with chaperone ASF1 and the replicative helicase subunit MCM2. *Protein Cell* **6**, 693–697 (2015). doi: 10.1007/s13238-015-0190-0; pmid: 26186914
38. N. Richet *et al*., Structural insight into how the human helicase subunit MCM2 may act as a histone chaperone together with ASF1 at the replication fork. *Nucleic Acids Res.* **43**, 1905–1917 (2015). doi: 10.1093/nar/gkv021; pmid: 25618846
39. H. Huang *et al*., A unique binding mode enables MCM2 to chaperone histones H3-H4 at replication forks. *Nat. Struct. Mol. Biol.* **22**, 618–626 (2015). doi: 10.1038/nsmb.3055; pmid: 26167883
40. A. Groth *et al*., Regulation of replication fork progression through histone supply and demand. *Science* **318**, 1928–1931 (2007). doi: 10.1126/science.1148992; pmid: 18096807
41. X. Ye *et al*., Defective S phase chromatin assembly causes DNA damage, activation of the S phase checkpoint, and S phase arrest. *Mol. Cell* **11**, 341–351 (2003). doi: 10.1016/S1097-2765(03)00037-6; pmid: 12620223
42. S. Ramachandran, S. Henikoff, MINCE-Seq: Mapping in vivo nascent chromatin with EdU and sequencing. *Methods Mol. Biol.* **1832**, 159–168 (2018). doi: 10.1007/978-1-4939-8663-7_8; pmid: 30073526
43. Z. Xu, J. Feng, Q. Li, Measuring genome-wide nascent nucleosome assembly using ReIN-Map. *Methods Mol. Biol.* **2196**, 117–141 (2021). doi: 10.1007/978-1-0716-0868-5_10; pmid: 32889717
44. A. Yesbolatova *et al*., The auxin-inducible degron 2 technology provides sharp degradation control in yeast, mammalian cells, and mice. *Nat. Commun.* **11**, 5701 (2020). doi: 10.1038/s41467-020-19532-z; pmid: 33177522
45. E. K. Monson, D. de Bruin, V. A. Zakian, The yeast Cac1 protein is required for the stable inheritance of transcriptionally repressed chromatin at telomeres. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 13081–13086 (1997). doi: 10.1073/pnas.94.24.13081; pmid: 9371803
46. F. Tessadori *et al*., Recurrent de novo missense variants across multiple histone H4 genes underlie a neurodevelopmental syndrome. *Am. J. Hum. Genet.* **109**, 750–758 (2022). doi: 10.1016/j.ajhg.2022.02.003; pmid: 35202563
47. S. Cheloufi *et al*., The histone chaperone CAF-1 safeguards somatic cell identity. *Nature* **528**, 218–224 (2015). doi: 10.1038/nature15749; pmid: 26659182
48. R. Franklin *et al*., Regulation of chromatin accessibility by the histone chaperone CAF-1 sustains lineage fidelity. *Nat. Commun.* **13**, 2350 (2022). doi: 10.1038/s41467-022-29730-6; pmid: 35487911
49. T. Ishiuchi *et al*., Early embryonic-like cells are induced by downregulating replication-dependent chromatin assembly. *Nat. Struct. Mol. Biol.* **22**, 662–671 (2015). doi: 10.1038/nsmb.3066; pmid: 26237512
50. L. Cheng *et al*., Chromatin Assembly Factor 1 (CAF-1) facilitates the establishment of facultative heterochromatin during pluripotency exit. *Nucleic Acids Res.* **47**, 11114–11131 (2019). doi: 10.1093/nar/gkz858; pmid: 31586391
51. C. M. Hammond, C. B. Strømme, H. Huang, D. J. Patel, A. Groth, Histone chaperone networks shaping chromatin function. *Nat. Rev. Mol. Cell Biol.* **18**, 141–158 (2017). doi: 10.1038/nrm.2016.159; pmid: 28053344
52. F. Mattiroli *et al*., DNA-mediated association of two histone-bound complexes of yeast Chromatin Assembly Factor-1 (CAF-1) drives

tetrasome assembly in the wake of DNA replication. *eLife* **6**, e22799 (2017). doi: 10.7554/eLife.22799; pmid: 28315523
53. K. Nozawa *et al*., Cryo-electron microscopy structure of the H3-H4 octasome: A nucleosome-like particle without histones H2A and H2B. *Proc. Natl. Acad. Sci. U.S.A.* **119**, e2206542119 (2022). doi: 10.1073/pnas.2206542119; pmid: 36322721
54. R. Vlijm *et al*., Nucleosome assembly dynamics involve spontaneous fluctuations in the handedness of tetrasomes. *Cell Rep.* **10**, 216–225 (2015). doi: 10.1016/j.celrep.2014.12.022; pmid: 25578730
55. C. Rouillon *et al*., CAF-1 deposits newly synthesized histones during DNA replication using distinct mechanisms on the leading and lagging strands. *Nucleic Acids Res.* **51**, 3770–3792 (2023). doi: 10.1093/nar/gkad171; pmid: 36942484
56. J. Fei *et al*., The prenucleosome, a stable conformational isomer of the nucleosome. *Genes Dev.* **29**, 2563–2575 (2015). doi: 10.1101/gad.272633.115; pmid: 26680301
57. A. Hamiche *et al*., Interaction of the histone (H3-H4)2 tetramer of the nucleosome with positively supercoiled DNA minicircles: Potential flipping of the protein from a left- to a right-handed superhelical form. *Proc. Natl. Acad. Sci. U.S.A.* **93**, 7588–7593 (1996). doi: 10.1073/pnas.93.15.7588; pmid: 8755519
58. T. Furuyama, S. Henikoff, Centromeric nucleosomes induce positive DNA supercoils. *Cell* **138**, 104–113 (2009). doi: 10.1016/j.cell.2009.04.049; pmid: 19596238
59. X. Ming *et al*., Kinetics and mechanisms of mitotic inheritance of DNA methylation and their roles in aging-associated methylome deterioration. *Cell Res.* **30**, 980–996 (2020). doi: 10.1038/s41422-020-0359-9; pmid: 32581343
60. Z. Otwinowski, W. Minor, Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997). doi: 10.1016/S0076-6879(97)76066-X; pmid: 27754618
61. A. J. McCoy *et al*., Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007). doi: 10.1107/S0021889807021206; pmid: 19461840
62. P. Emsley, K. Cowtan, Coot: Model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004). doi: 10.1107/S0907444904019158; pmid: 15572765
63. P. D. Adams *et al*., PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010). doi: 10.1107/S0907444909052925; pmid: 20124702
64. V. B. Chen *et al*., MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 12–21 (2010). doi: 10.1107/S0907444909042073; pmid: 20057044
65. P. N. Dyer *et al*., Reconstitution of nucleosome core particles from recombinant histones and DNA. *Methods Enzymol.* **375**, 23–44 (2004). doi: 10.1016/S0076-6879(03)75002-2; pmid: 14870657
66. P. T. Lowary, J. Widom, New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. *J. Mol. Biol.* **276**, 19–42 (1998). doi: 10.1006/jmbi.1997.1494; pmid: 9514715
67. B. Kastner *et al*., GraFix: Sample preparation for single-particle electron cryomicroscopy. *Nat. Methods* **5**, 53–55 (2008). doi: 10.1038/nmeth1139; pmid: 18157137
68. A. Punjani, J. L. Rubinstein, D. J. Fleet, M. A. Brubaker, cryoSPARC: Algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017). doi: 10.1038/nmeth.4169; pmid: 28165473
69. K. Zhang, Gctf: Real-time CTF determination and correction. *J. Struct. Biol.* **193**, 1–12 (2016). doi: 10.1016/j.jsb.2015.11.003; pmid: 26592709
70. J. Zivanov *et al*., New tools for automated high-resolution cryo-EM structure determination in RELION-3. *eLife* **7**, e42166 (2018). doi: 10.7554/eLife.42166; pmid: 30412051
71. D. Asarnow, E. Palovcak, Y. Cheng, UCSF pyem v0.5. *Zenodo* (2019). doi: 10.5281/zenodo.3576630
72. E. F. Pettersen *et al*., UCSF Chimera—A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004). doi: 10.1002/jcc.20084; pmid: 15264254
73. T. D. Goddard *et al*., UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci.* **27**, 14–25 (2018). doi: 10.1002/pro.3235; pmid: 28710774
74. A. Osakabe *et al*., Nucleosome formation activity of human somatic nuclear autoantigenic sperm protein (sNASP). *J. Biol. Chem.* **285**, 11913–11921 (2010). doi: 10.1074/jbc.M109.083238; pmid: 20167597
75. J. Xiong *et al*., Cooperative action between SALL4A and TET proteins in stepwise oxidation of 5-methylcytosine. *Mol. Cell* **64**, 913–925 (2016). doi: 10.1016/j.molcel.2016.10.013; pmid: 27840027
76. S. Ramachandran, S. Henikoff, Transcriptional regulators compete with nucleosomes post-replication. *Cell* **165**, 580–592 (2016). doi: 10.1016/j.cell.2016.02.062; pmid: 27062929
77. S. Liu *et al*., RPA binds histone H3-H4 and functions in DNA replication-coupled nucleosome assembly. *Science* **355**, 415–420 (2017). doi: 10.1126/science.aah4712; pmid: 28126821
78. J. Yu *et al*., Analysis of local chromatin states reveals gene transcription potential during mouse neural progenitor cell differentiation. *Cell Rep.* **32**, 107953 (2020). doi: 10.1016/j.celrep.2020.107953; pmid: 32726618
79. B. Langmead, S. L. Salzberg, Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012). doi: 10.1038/nmeth.1923; pmid: 22388286
80. K. Chen *et al*., DANPOS: Dynamic analysis of nucleosome position and occupancy by sequencing. *Genome Res.* **23**, 341–351 (2013). doi: 10.1101/gr.142067.112; pmid: 23193179
81. J. H. I. Haarhuis *et al*., A Mediator-cohesin axis controls heterochromatin domain formation. *Nat. Commun.* **13**, 754 (2022). doi: 10.1038/s41467-022-28377-7; pmid: 35136067
82. J. Lipfert, M. Wiggin, J. W. Kerssemakers, F. Pedaci, N. H. Dekker, Freely orbiting magnetic tweezers to directly monitor changes in the twist of nucleic acids. *Nat. Commun.* **2**, 439 (2011). doi: 10.1038/ncomms1450; pmid: 21863006
83. P. Chen *et al*., Functions of FACT in breaking the nucleosome and maintaining its integrity at the single-nucleosome level. *Mol. Cell* **71**, 284–293.e4 (2018). doi: 10.1016/j.molcel.2018.06.020; pmid: 30029006
84. A. J. te Velthuis, J. W. Kerssemakers, J. Lipfert, N. H. Dekker, Quantitative guidelines for force calibration through spectral analysis of magnetic tweezers data. *Biophys. J.* **99**, 1292–1302 (2010). doi: 10.1016/j.bpj.2010.06.008; pmid: 20713015
85. Z. Li, Code for analysis and plotting data from freely orbiting magnetic tweezers (1.0). *Zenodo* (2023); https://doi.org/10.5281/zenodo.8166291.

ACKNOWLEDGMENTS

We thank B. Zhu, X. Huang, and staff members at the Center for Biological Imaging (CBI) of the Institute of Biophysics (IBP), Chinese Academy of Sciences, for support in cryo-EM data collection; X. Yu at IBP for technical support with SEC-MALS analysis; and staff scientists at Shanghai Synchrotron Radiation Facility (SSRF) beamlines BL17U and BL19U for assistance in x-ray data collection. We also thank M. Jin and Y. Cong at Shanghai Institute of Biochemistry and Cell Biology and M. Zhou, Y. Wang, P. Zhu, and D. Cao at IBP for participation during the early stages of this study. **Funding:** This work was funded by Ministry of Science and Technology of China grant 2019YFA0508900 (R.-M.X.); National Natural Science Foundation of China grants 31991162, 91853204, and 92153302 (R.-M.X.) and 32288102 (B.Z.); Ministry of Science and Technology of China grants 2018YFE0203300 and 2017YFA0506600 (C.-P.L.); Chinese Academy of Sciences Strategic Priority Research Program XDB37010100 (R.-M.X.) and XDB39010100 (B.Z.); Chinese Academy of Sciences Youth Innovation Promotion Association grants 2018125 (C.-P.L.), 2017131 (Z.Y.), 2020097 (J.X.), and 2017133 (Z.Q.Z.); and the New Cornerstone Investigator Program (B.Z.). **Author contributions:** Conceptualization: R.-M.X., C.-P.L., Z.G.Z.; Methodology: C.-P.L., C.Y., M.W., W.L., Z.Q.Z., X.Z., B.Z., G.L., R.-M.X.; Investigation: C.-P.L., Z.Y., J.X., J.H., A.S., D.D., J.Y., P.H., K.Z., Z.L.; Visualization: C.-P.L., C.Y., M.W.; Funding acquisition: R.-M.X., C.-P.L.; Project administration: R.-M.X., N.Y.; Supervision: R.-M.X., G.L., B.Z., N.Y., W.L.; Writing – original draft: C.-P.L., R.-M.X.; Writing – review and editing: All authors. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data and materials used in this study are available for academic purposes upon request. Atomic coordinates and associated structure factors from crystallographic studies have been deposited in the Protein Data Bank (PDB) under accession numbers 7Y5K, 7Y5L, and 7Y5O for the models belonging to the $C222_1$, C2, and $P2_1$ space groups, respectively. EM maps and fitted structural models have been deposited in the Electron Microscopy Data Bank (EMDB) and PDB with the following accession information: the 1:1 CAF1LC-H3-H4 complex, EMD-33625 and 7Y5U; the 2:2 CAF1LC-H3-H4 complex, EMD-33626 and 7Y5V; the monomeric CAF-1-H3-H4 complex, EMD-35661 and 8IQG; the dimeric CAF-1-H3-H4 complex, EMD-35660 (composite map), EMD-35708 (consensus map), EMD-35709 (local map for protomer I), EMD-35710 (local map for protomer II), and 8IQF; left-handed di-tetrasome, EMD-33627 and 7Y5W; CAF1LC bound right-handed ditetrasome, EMD-33630 and 7Y60; right-handed ditetrasome bound by two CAF1LC complexes, EMD-33631 and 7Y61; the single CAF-1–bound right-handed ditetrasome, EMD-36013 and 8J6S; and the double CAF-1–bound right-handed ditetrasome, EMD-36014 and 8J6T. The raw sequence data reported in this paper have been deposited in the Genome Sequence Archive in the National Genomics Data Center, China National Center for Bioinformation, Chinese Academy of Sciences (GSA-Human: HRA002511), and are publicly accessible at https://ngdc.cncb.ac.cn/gsa-human. The code for analyzing and plotting FOMT data is available for downloading via Zenodo (*85*). **License information:** Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.add8673
Figs. S1 to S14
Tables S1 to S3
MDAR Reproducibility Checklist
Movie S1

Submitted 9 July 2022; resubmitted 19 May 2023
Accepted 19 July 2023
10.1126/science.add8673

RESEARCH ARTICLE SUMMARY

CANCER GENETICS

Oncogene-like addiction to aneuploidy in human cancers

Vishruth Girish†, Asad A. Lakhani†, Sarah L. Thompson‡, Christine M. Scaduto‡, Leanne M. Brown, Ryan A. Hagenson, Erin L. Sausville, Brianna E. Mendelson, Pranav K. Kandikuppa, Devon A. Lukow, Monet Lou Yuan, Eric C. Stevens, Sophia N. Lee, Klaske M. Schukken, Saron M. Akalu, Anand Vasudevan, Charles Zou, Barbora Salovska, Wenxue Li, Joan C. Smith, Alison M. Taylor, Robert A. Martienssen, Yansheng Liu, Ruping Sun, Jason M. Sheltzer*

INTRODUCTION: It has been known for more than 100 years that human cancers exhibit pervasive aneuploidy, or chromosome copy number changes. For instance, about 25% of cancers exhibit gains of the q arm of chromosome 1. However, despite the prevalence of aneuploidy across cancer types, its role in tumorigenesis has remained poorly defined. Our ability to uncover the function of these large-scale copy number alterations has been hampered by our inability to experimentally manipulate chromosome dosage in cancer. Nonetheless, as aneuploidy is common across malignancies but rare in normal tissue, drugs that exhibit selective toxicity toward aneuploid cells could be useful anticancer agents.

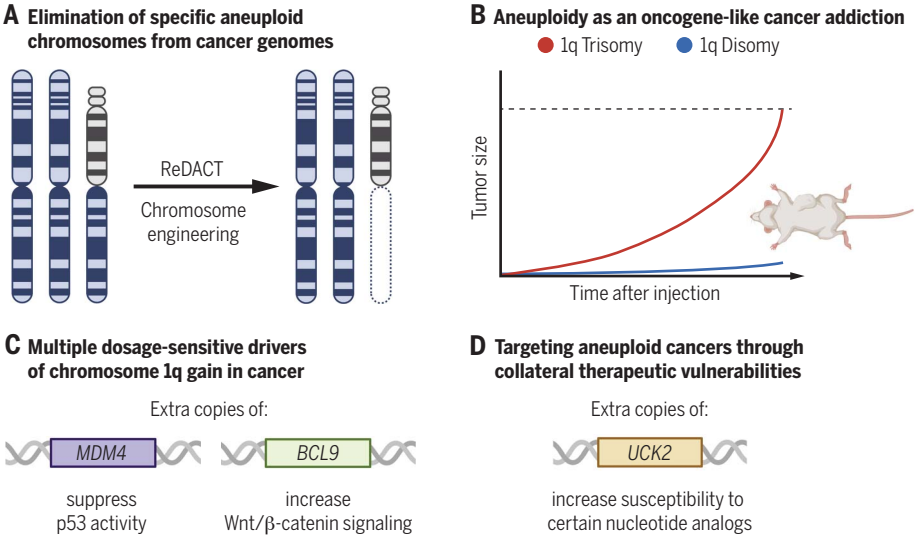
RATIONALE: Although aneuploidies have resisted close analysis, previous research has led to the discovery of a phenomenon called "oncogene addiction." An oncogene-addicted cancer is dependent on the expression of an individual oncogene for continued malignant growth, and loss or inhibition of that oncogene is sufficient to induce cancer regression. As specific aneuploidies such as the gain of chromosome 1q are frequent events in diverse cancer types, we hypothesized that certain aneuploidies could themselves represent oncogene-like cancer addictions. To test this hypothesis, we developed ReDACT (Restoring Disomy in Aneuploid cells using CRISPR Targeting), a set of chromosome engineering tools that allow us to eliminate individual aneuploid chromosomes from cancer genomes. Using ReDACT, we created and then characterized a panel of isogenic cells that have or lack common cancer aneuploidies.

RESULTS: We found that eliminating the trisomy of chromosome 1q from cancer cell lines harboring this alteration almost completely abolished anchorage-independent growth and xenograft formation. Similarly, eliminating the 1q trisomy from a nonmalignant cell line blocked *RAS*-mediated transformation. Prolonged growth in vitro or in vivo after aneuploidy elimination in cancer cell lines led to karyotype evolution, and 1q-disomic cells were eventually outcompeted by cells that had recovered the 1q trisomy. In contrast, removing other trisomic chromosomes from cancer cells had variable effects on malignant growth, demonstrating that different aneuploidies have distinct phenotypic consequences for cancer development.

An analysis of clinical sequencing data demonstrated that chromosome 1q gains arise early during tumorigenesis and are mutually exclusive with mutations in the tumor suppressor *TP53*, suggesting that 1q trisomies could represent a mutation-independent mechanism for blocking p53 signaling. Consistent with this, we demonstrated that ReDACT-mediated elimination of chromosome 1q trisomies increased the expression of p53 target genes in *TP53* wild-type cell lines. We traced this suppression of p53 function to the triplication of *MDM4*, a p53 inhibitor encoded on chromosome 1q, and we found that deleting a single copy of *MDM4* impaired the growth of 1q-trisomic cells, whereas moderate overexpression of *MDM4* rescued the growth of 1q-disomic cells.

Finally, we demonstrated that chromosome 1q gains result in the overexpression of *UCK2*, a nucleotide kinase encoded on chromosome 1q that is also required for the cytotoxicity of certain anticancer nucleotide analogs. We determined that several different 1q-trisomic cell lines displayed enhanced sensitivity to these compounds owing to the up-regulation of *UCK2*, revealing that 1q aneuploidy can also represent a tractable cancer vulnerability.

CONCLUSION: Certain aneuploidies that are commonly found in tumor genomes play a central role in cancer development, and eliminating these aneuploidies compromises malignant growth potential. At the same time, aneuploidy causes collateral therapeutic vulnerabilities that can be targeted to selectively eliminate cells with chromosome dosage imbalances. The development of flexible chromosome engineering methodologies like ReDACT will enable additional experiments to further unravel the consequences of aneuploidy in development and disease. ■



Chromosomal engineering to investigate the effects of aneuploidy. (**A**) ReDACT enables the targeted deletion of aneuploid chromosomes. (**B**) Loss of an extra copy of chromosome 1q compromises malignant growth. (**C**) *MDM4* and *BCL9* are dosage-sensitive drivers of chromosome 1q gain in cancer. (**D**) Chromosome 1q gain can be targeted therapeutically with UCK2-specific nucleotide analogs.

***Corresponding author. Email: jason.sheltzer@yale.edu**
†These authors contributed equally to this work.
‡These authors contributed equally to this work.
Cite this article as V. Girish *et al.*, *Science* 381, eadg4521 (2023). DOI: 10.1126/science.adg4521

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.adg4521

RESEARCH ARTICLE

CANCER GENETICS

Oncogene-like addiction to aneuploidy in human cancers

Vishruth Girish[1,2]†, Asad A. Lakhani[3]†, Sarah L. Thompson[1]‡, Christine M. Scaduto[3]‡, Leanne M. Brown[1], Ryan A. Hagenson[1], Erin L. Sausville[3], Brianna E. Mendelson[1], Pranav K. Kandikuppa[1], Devon A. Lukow[1], Monet Lou Yuan[3], Eric C. Stevens[1], Sophia N. Lee[1], Klaske M. Schukken[1], Saron M. Akalu[1], Anand Vasudevan[1], Charles Zou[1], Barbora Salovska[1], Wenxue Li[1], Joan C. Smith[1], Alison M. Taylor[4], Robert A. Martienssen[3,5], Yansheng Liu[1], Ruping Sun[6], Jason M. Sheltzer[1]*

Most cancers exhibit aneuploidy, but its functional significance in tumor development is controversial. Here, we describe ReDACT (Restoring Disomy in Aneuploid cells using CRISPR Targeting), a set of chromosome engineering tools that allow us to eliminate specific aneuploidies from cancer genomes. Using ReDACT, we created a panel of isogenic cells that have or lack common aneuploidies, and we demonstrate that trisomy of chromosome 1q is required for malignant growth in cancers harboring this alteration. Mechanistically, gaining chromosome 1q increases the expression of *MDM4* and suppresses p53 signaling, and we show that *TP53* mutations are mutually exclusive with 1q aneuploidy in human cancers. Thus, tumor cells can be dependent on specific aneuploidies, raising the possibility that these "aneuploidy addictions" could be targeted as a therapeutic strategy.

Chromosome copy number changes, otherwise known as aneuploidy, are a ubiquitous feature of tumor genomes (*1*, *2*). While the pervasiveness of aneuploidy in cancer has been known for more than a century (*3*, *4*), the role of aneuploidy in tumor development has remained controversial (*5*–*8*). Chromosome gains have been proposed to serve as a mechanism for increasing the dosage of tumor-promoting genes that are found within altered regions (*9*, *10*). However, proof of this hypothesis is lacking, and it has alternately been suggested that aneuploidy could arise as a result of the loss of checkpoint control that frequently occurs in advanced malignancies (*11*). Indeed, individuals with Down syndrome, which is caused by the triplication of chromosome 21, have a markedly decreased risk of developing most solid cancers, suggesting that, in certain cases, aneuploidy may actually have tumor-suppressive properties (*12*).

Our ability to directly investigate the role of aneuploidy in cancer has historically been limited by the experimental difficulties involved in manipulating entire chromosome arms. Over the past 40 years, cancer researchers have used the standard tools of molecular genetics, including gene overexpression, knockdown, and mutagenesis, to develop a deep understanding of many individual oncogenes and tumor suppressors (*13*, *14*). For instance, the biological functions of genes such as *KRAS* and *TP53* were elucidated in part by creating and analyzing isogenic cell lines that express or lack these genes (*15*, *16*). However, existing approaches for single-gene manipulations are insufficient to interrogate the chromosome-scale changes that affect hundreds of genes simultaneously. The consequences of eliminating specific aneuploid chromosomes from human cancer cells have not previously been established.

Studies of individual cancer driver genes led to the discovery of a phenomenon called "oncogene addiction," in which loss or inhibition of a single oncogene is sufficient to induce cancer regression (*17*). For example, mutations in *KRAS* cause the development of pancreas cancer, and genetically ablating *KRAS* in a "*KRAS*-addicted" pancreas tumor blocks growth and triggers apoptosis (*18*). Previous cancer genome sequencing projects have revealed that the aneuploidy patterns observed in human tumors are nonrandom, and specific events such as the gain of chromosome 1q or 8q occur more often than expected by chance (*1*, *19*). We speculated that these recurrent aneuploidies could themselves represent a type of cancer "addiction," analogous to the concept of oncogene addictions. To investigate this hypothesis, we developed a set of computational and functional techniques to facilitate the analysis of cancer aneuploidy.

[1]Yale University School of Medicine, New Haven, CT 06511, USA. [2]Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA. [3]Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724, USA. [4]Columbia University School of Medicine, New York, NY 10032, USA. [5]Howard Hughes Medical Institute, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724, USA. [6]Department of Laboratory Medicine and Pathology, University of Minnesota, Minneapolis, MN 55455, USA.
*Corresponding author. Email: jason.sheltzer@yale.edu
†These authors contributed equally to this work.
‡These authors contributed equally to this work.

Results

Specific chromosome gains recurrently occur early in cancer development

We recently established a computational approach to leverage multisample tumor sequencing data to determine the relative timing of somatic copy number alterations in cancer evolution (*20*). We applied this tool to investigate the timing of aneuploidy events in a cohort of patients with breast cancer (BRCA) or melanoma (MEL) (*21*, *22*). We modeled the relative timing of somatic copy number alterations in whole-genome sequences from these tumor samples by assuming that (i) somatic point mutations accumulate over time at a rate that is proportional to DNA copy number and (ii) the multiplicity of early point mutations increases with copy number gains. We found that specific chromosome copy number changes are consistently observed early in tumor development (Fig. 1, A and B). Notably, we observed that chromosome 1q gains are recurrently the first copy number alterations that occur in breast cancer evolution, and these gains are also among the first alterations in melanoma evolution. In general, we found that common aneuploidies arose earlier in tumor development than less-common aneuploidies, in agreement with the assumption that early somatic alterations are likely to be fitness-driving events (Fig. 1C) (*23*). However, the correlation between frequency and timing was not maintained across all chromosomes. For instance, in breast cancer, chromosome 8q gains and chromosome 1q gains occurred with similar frequencies, but we found that 1q gains consistently arose earlier during tumor development than 8q gains. We conclude that, as previously observed with oncogenic point mutations (*24*), specific chromosome gains occur in a defined temporal order, and we speculate that aneuploidies that are consistently gained early during tumorigenesis may enhance cancer fitness.

Specific chromosome gains are associated with altered mutational patterns and cancer progression

In instances where two oncogenes converge to activate the same pathway, cancers frequently acquire mutations in either gene but not both (*25*). If chromosome gains play an oncogene-like role supporting cancer growth, then specific aneuploidies may also be expected to exhibit mutual exclusivity with individual oncogenic mutations. To investigate this possibility, we calculated patterns of mutual exclusivity between chromosome arm gains and mutations across 23,544 cancer patients (*26*, *27*). We detected several hundred instances in which aneuploidies and mutations co-occur less often than expected by chance both within individual cancer types and in a pan-cancer analysis (Fig. 1, D and E; fig. S1; and

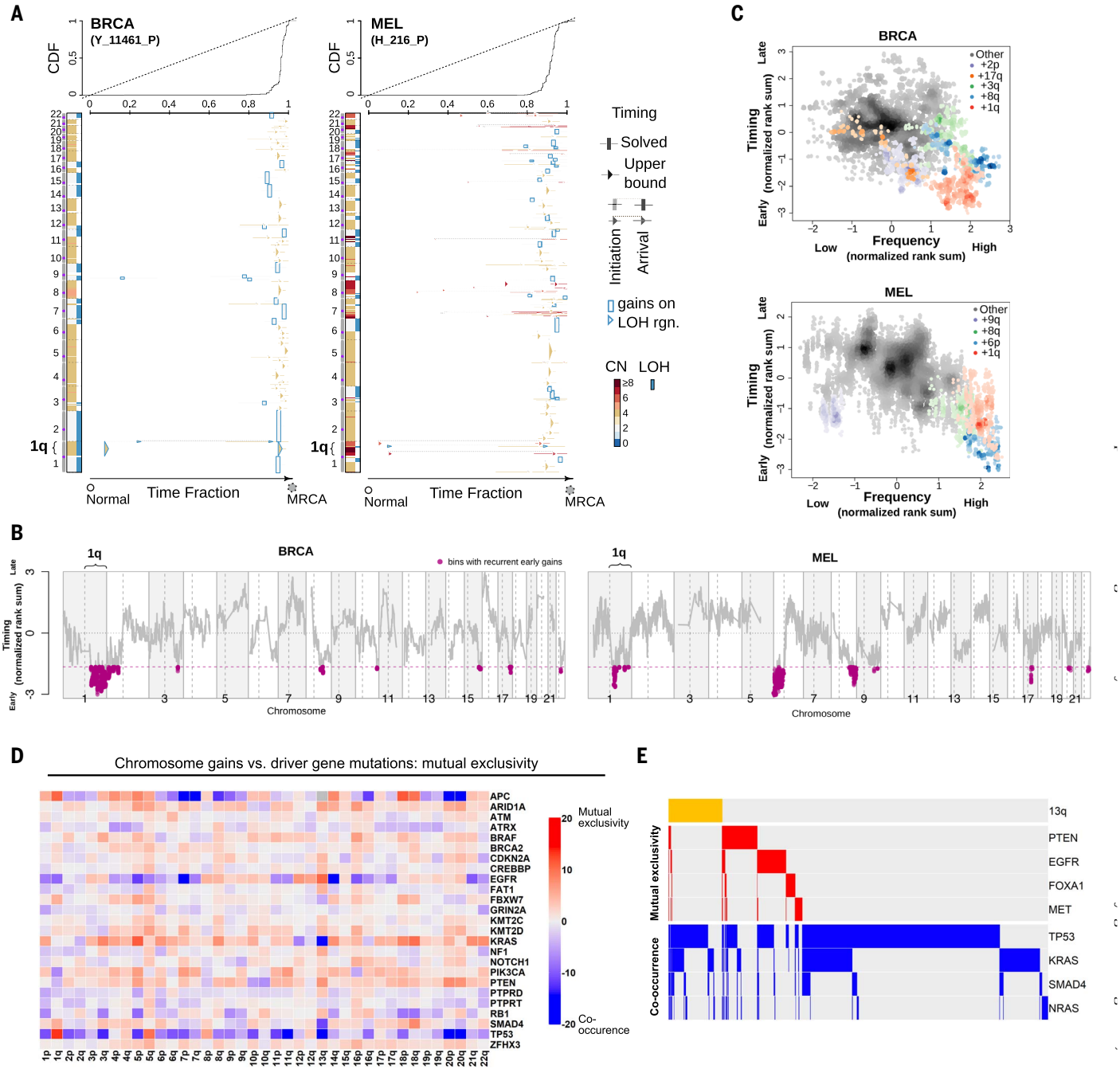


Fig. 1. Specific chromosome gains arise early in tumor development and are mutually exclusive with driver gene mutations. (**A**) The inferred timing of somatic copy number gains in the evolution of two tumors. A breast tumor is shown on the left and a melanoma on the right. Copy number (CN) states along the genome are shown on the left in each panel and color coded. The plot visualizes the time fraction of somatic evolution from germline to the most recent common ancestor (MRCA) of the patient tumor sample. For each copy number segment, the inferred timing is shown as a rectangle (exactly solved timing) or an arrow (upper bounds of timing when the timing solutions are not unique) with the same color-coding as its CN. The top panel shows the cumulative distribution (CDF) of the timing. Genome doubling (GD) can be observed as the punctuated gains occurring in a narrow time window, and chromosome 1q gains appear to be extremely early and preceding GD in these two tumors. (**B**) Recurrent early gains of chromosome 1q in BRCA (n = 38 tumor samples) and MEL (n = 37 tumor samples). For each tumor type, we converted the timing of gains into ranks for genomic bins within a patient and computed the rank sums across patients for each bin. The normalized rank sums for each genomic bin are shown for BRCA and MEL. The large negative values indicate recurrent early initiating gains. We used the normalized rank sums to test against the null hypothesis (no regions show recurrent early gain across patients). Bins from chromosome 1q reject this null for both tumor types (with 90% confidence level). (**C**) The timing of a gain compared with the frequency of its occurrence in BRCA (n = 38 tumor samples) and MEL (n = 37 tumor samples). The points on the plots show the timing of gain of a genomic bin versus its frequency of copy number gain. Colors represent chromosomal arms, and color saturation indicates the density of points. Both the timings and frequencies were transformed into normalized rank sums (see methods). In total, 15 out of 21 BRCA patients and 24 out of 37 MEL patients exhibited arm-scale gains of chromosome 1q. (**D**) A pan-cancer analysis of mutual exclusivity between mutations in 25 commonly mutated cancer genes and chromosome arm gain events. The complete results of this analysis are included in table S1. (**E**) Mutual exclusivity and co-occurrence patterns between one representative chromosome gain (+13q, orange bars at the top), and point mutations in several different cancer driver genes.

table S1). For instance, *KRAS* mutations are mutually exclusive with chromosome 18q gains in pancreatic cancer, whereas *BRAF* mutations are mutually exclusive with chromosome 20q gains in colorectal cancer (fig. S1 and table S1). These results are consistent with our hypothesis that specific chromosome gains can play an oncogene-like role in cancer, thereby making the acquisition of certain oncogenic mutations redundant in the presence of that aneuploidy.

High levels of aneuploidy are generally associated with poor outcomes for cancer patients (*28*–*30*). However, it is less clear whether specific copy number changes drive tumor progression, or whether the aneuploid state itself represents a universal risk factor. We calculated the association between patient outcome and copy number gains affecting every chromosome band across 10,884 patients and 33 cancer types from The Cancer Genome Atlas (TCGA) (*31*). We discovered that certain copy number alterations were commonly prognostic across multiple cancer types, particularly gains affecting chromosome 1q (fig. S2, A to C, and table S2A). The strong association between 1q gains and disease progression was robust to the inclusion of multiple clinical variables, including patient age, sex, tumor stage, and tumor grade (fig. S2D and table S2B). 1q copy number correlated with hallmarks of aggressive disease in genetically diverse cancer types, including with Gleason score in prostate adenocarcinoma and with thrombocytopenia in acute myeloid leukemia (fig. S2E). We performed a similar analysis for cancer-associated mutations, and we found that the only gene for which mutations were prognostic in more than four cancer types was *TP53* (fig. S2, A and D). These results illustrate that specific chromosome gain events, particularly involving regions of chromosome 1q, are robust pan-cancer markers for the risk of disease progression.

Loss of trisomy 1q blocks malignant growth in human cancers

The computational analyses described above highlight several similarities between chromosome copy number gains and driver mutations, raising the possibility that these aneuploidies could represent oncogene-like cancer addictions. The oncogene addiction paradigm was first established by developing genetic techniques to eliminate individual genes from established cancer cell lines (*16*, *17*). To conduct comparable assays with aneuploidy, we created a set of approaches collectively called ReDACT (Restoring Disomy in Aneuploid cells using CRISPR Targeting) (Fig. 2A). In the first approach, called ReDACT-NS (Negative Selection), we integrate a copy of herpesvirus thymidine kinase (*HSV-TK*) onto an aneuploid chromosome of interest. Then, the cells are transfected with a guide RNA (gRNA) that cuts between the integrant and the centromere and are treated with ganciclovir, which is toxic to cells that express *HSV-TK* (*32*). Loss of the aneuploid chromosome harboring *HSV-TK* allows cells to survive ganciclovir selection. In the second approach, called ReDACT-TR (Telomere Replacement), cells are cotransfected with a gRNA that cuts near the centromere of an aneuploid chromosome along with a cassette encoding ~100 repeats of the human telomere seed sequence (*33*). CRISPR cleavage coupled with integration of the telomeric seed sequence leads to loss of an aneuploid chromosome arm and formation of a de novo telomere. In the third approach, called ReDACT-CO (CRISPR Only), we took advantage of prior reports demonstrating that, in rare circumstances, CRISPR cleavage by itself is sufficient to trigger chromosome loss (*34*, *35*), and we transfected cells with a gRNA targeting an aneuploid chromosome arm without any other selection markers. We successfully applied all three approaches to create clones derived from human cell lines that had lost specific aneuploid chromosomes.

We first focused on aneuploidies of chromosome 1q, as we found that 1q gains were an early event in multiple cancer types and were strongly associated with disease progression (Fig. 1 and fig. S2). We targeted the 1q trisomy in the A2058 melanoma cell line, AGS gastric cancer cell line, and A2780 ovarian cancer cell line. We generated multiple independent derivatives of each line in which a single copy of chromosome 1q had been eliminated, thereby producing cell lines that were disomic rather than trisomic for chromosome 1q. We verified loss of the 1q trisomy using short multiply aggregated sequence homologies sequencing (SMASH-Seq), an approach to determine DNA copy number (*36*), and by G-banding analysis of metaphase spreads (Fig. 2B, figs. S3 and S4, and table S3). Loss of the 1q trisomy decreased the expression of genes encoded on chromosome 1q by an average of 26% at the RNA level and 21% at the protein level (Fig. 2C). These results suggest that chromosome loss causes a substantial down-regulation of genes encoded on an aneuploid chromosome, although these effects can be buffered to some extent by cellular dosage compensation (*37*).

Next, we tested whether losing the 1q trisomy affects malignant growth in cancer cells. Toward that end, we quantified anchorage-independent colony formation, an in vitro proxy for malignant potential (*38*), in the 1q-trisomic and 1q-disomic cells. While 1q-trisomic A2058, A2780, and AGS cells displayed robust colony formation, multiple independent 1q-disomic clones derived from each cell line exhibited minimal anchorage-independent growth (Fig. 2D). We then performed contralateral subcutaneous injections with each cell line to test whether aneuploidy loss affected tumor growth in vivo. Consistent with our colony formation assays, we observed that 1q-trisomic A2058 and A2780 cells rapidly formed large tumors, whereas 1q-disomic cells displayed minimal tumor growth (Fig. 2E). At the end of these assays, the trisomic cells had formed tumors that were, on average, 25-fold larger than the tumors formed by the 1q-disomic cells. For the AGS cancer cell line, neither the trisomic nor the disomic cells formed tumors after subcutaneous injection (fig. S5). Finally, we performed proliferation assays to measure the doubling time of the 1q-trisomic and the 1q-disomic cells in culture (fig. S6, A to C). The aneuploidy-loss clones divided more slowly in vitro compared with the 1q-trisomic cells, although the difference in doubling time (~35%) was substantially less than the differences observed in the soft agar and xenograft assays. In total, these results suggest that multiple human cancer cell lines are dependent on the presence of a third copy of chromosome 1q to support malignant growth and that elimination of this aneuploid chromosome compromises their tumorigenic potential. Furthermore, we note that this phenotypic pattern, in which aneuploidy loss causes a moderate effect on in vitro doubling but a severe effect on anchorage-independent growth and xenograft formation, resembles the previously reported consequences of eliminating bona fide oncogene addictions (*16*, *39*).

Loss of trisomy 1q prevents malignant transformation

We discovered that 1q gains were commonly the first copy number alteration to occur during breast tumor development (Fig. 1, A to C). We therefore hypothesized that, in addition to being required for cancer growth, aneuploidy of chromosome 1q may directly promote cellular transformation. To test this, we performed chromosome engineering in MCF10A, an immortal but nontumorigenic mammary epithelial cell line. SMASH-Seq revealed that this cell line harbors a trisomy of chromosome 1q, and we successfully applied ReDACT-CO and ReDACT-TR to generate derivatives of MCF10A with two rather than three copies of 1q (Fig. 2F, fig. S3D, and table S3). We then attempted to transform the 1q-trisomic and 1q-disomic cells by transducing them with a retrovirus encoding the $HRAS^{G12V}$ oncogene. $HRAS^{G12V}$ expression was sufficient to transform trisomic MCF10A, as these cells were able to form colonies in soft agar and grow as xenografts in nude mice (Fig. 2, G and H). In contrast, 1q-disomic MCF10A clones expressing $HRAS^{G12V}$ exhibited impaired colony formation and were unable to produce tumors in vivo, demonstrating that loss of the trisomic chromosome prevented cellular transformation. These results

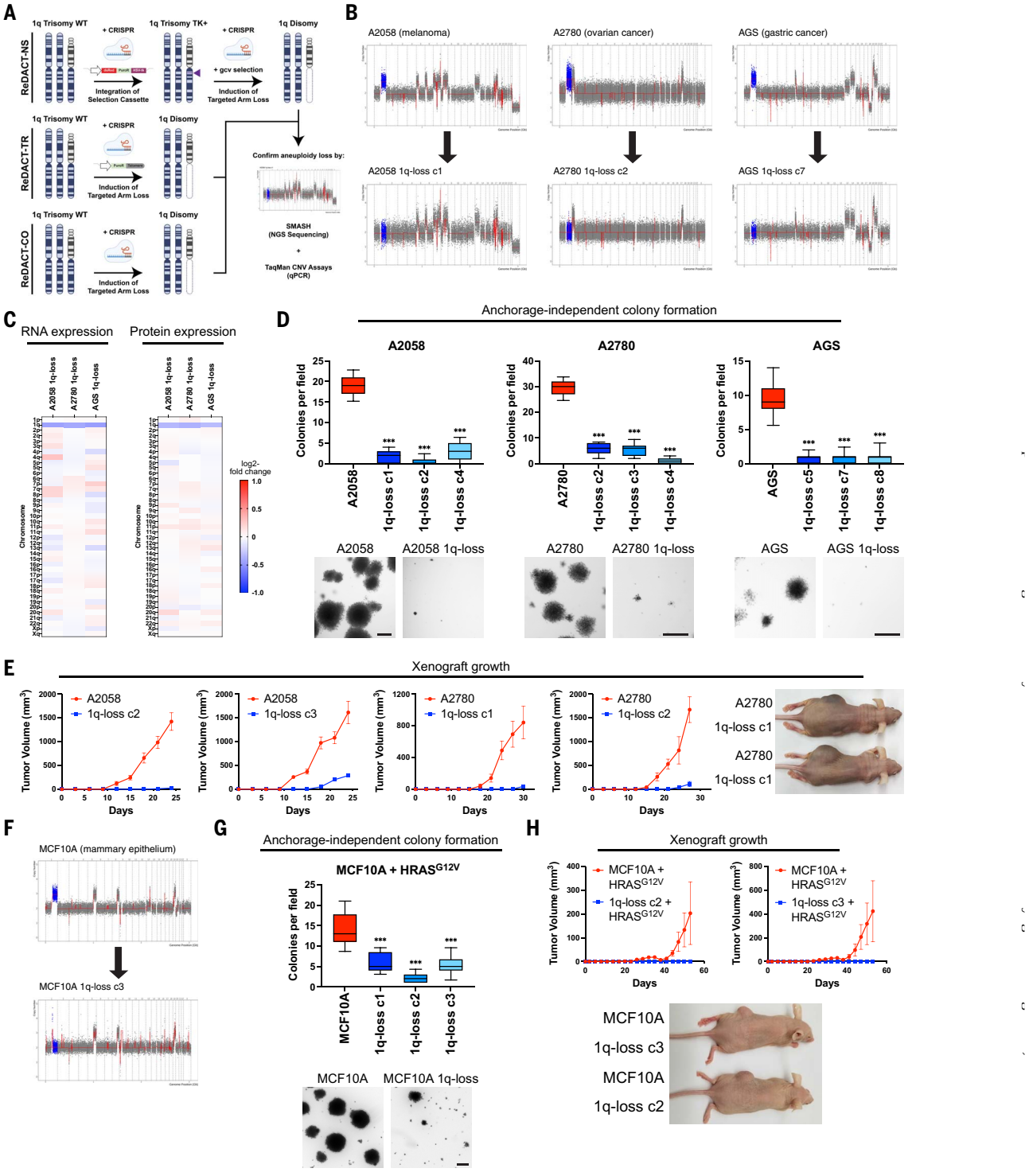


Fig. 2. Phenotypic effects of losing chromosome 1q aneuploidy.
(**A**) Chromosomal engineering strategies for the targeted deletion of chromosome arms: (i) ReDACT-NS: Using CRISPR-Cas9 homology-directed repair, we integrated a positive-negative selection cassette encoding a fluorescent reporter, a positive selection marker, and a negative selection marker (HSV thymidine kinase) at a centromere-proximal region on chromosome 1q. We induced arm loss by generating a double-stranded DNA (dsDNA) break centromere-proximal to the cassette with Cas9 and isolated clonal populations of cells that were ganciclovir-resistant. (ii) ReDACT-TR: We induced arm loss by generating a dsDNA break at a centromere-proximal location with Cas9 while providing cells with an ectopic telomere seed sequence for repair. (iii) ReDACT-CO: We induced arm loss by generating a dsDNA break at a centromere-proximal location with Cas9 and isolated clonal populations of cells. For all three approaches, we screened clonal populations of cells for targeted chromosome loss through TaqMan copy number assays and validated their karyotypes through SMASH sequencing. (**B**) Representative SMASH karyotypes of the 1q-disomic clones generated from the 1q-trisomic cancer cell lines A2780, AGS, and A2058. Chromosome 1q is highlighted in blue. A complete list of aneuploidy-loss clones and how they were generated is included in table S3. (**C**) 1q-disomic clones display decreased RNA expression and protein expression

of genes encoded on chromosome 1q. RNA expression data were obtained through bulk RNA-seq and represent the average expression of genes by chromosome arm across multiple 1q-disomic clones for each cell line. Protein expression data were obtained through mass spectrometry, and representative data from one 1q-disomic clone are shown for each cell line. Data are log2 transformed, normalized to the parental cell line, and adjusted so that the mean expression across all chromosomes is 0. (**D**) 1q-disomic clones exhibit decreased anchorage-independent growth. The micrographs display representative images of colony formation for 1q-trisomic and 1q-disomic clones. (**E**) 1q-disomic clones exhibit impaired xenograft growth in vivo. 1q-trisomic and 1q-disomic cells were injected contralaterally and subcutaneously into immunocompromised mice. The graphs display the mean ± SEM for each trial. Representative mice are shown on the right. (**F**) SMASH karyotype of a 1q-disomic clone generated from the mammary epithelial cell line MCF10A. Chromosome 1q is highlighted in blue. (**G**) 1q-disomic MCF10A clones transduced with $HRAS^{G12V}$ exhibit decreased anchorage-independent growth relative to 1q-trisomic MCF10A cells. (**H**) 1q disomic MCF10A clones transduced with $HRAS^{G12V}$ clones exhibit impaired xenograft growth in vivo. 1q-trisomic and 1q-disomic cells were injected contralaterally and subcutaneously into immunocompromised mice. The graphs display the mean ± SEM for each trial. Representative mice are shown below. For anchorage-independent growth assays in (D) and (G), the boxplots represent the 25th, 50th, and 75th percentiles of colonies per field, while the whiskers represent the 10th and 90th percentiles. Unpaired *t* test, $n = 15$ fields of view, data from representative trial ($n \geq 2$ total trials). Representative images are shown below. Scale bars, 250 μm. $^{**}P < 0.005$, $^{***}P < 0.0005$.

are consistent with our finding that 1q gains are an early event during breast cancer development and demonstrate that specific aneuploidies can cooperate with oncogenes to transform nonmalignant cells.

Robust anchorage-independent growth in human cancer cell lines subjected to CRISPR cutting and ganciclovir selection

To confirm that our findings were a specific consequence of aneuploidy loss, we generated and tested a series of control clones subjected to various CRISPR manipulations that did not induce loss of the 1q trisomy (figs. S7 and S8A). These control clones included the following: (i) Cell lines harboring a CRISPR-mediated integration of the *HSV-TK* cassette that were not subjected to selection for 1q loss. (ii) Cell lines in which the *HSV-TK* cassette was deleted using two gRNAs followed by ganciclovir selection. (iii) Cell lines transfected with a 1q-targeting gRNA in which the lesion was repaired without inducing chromosome loss. (iv) Cell lines transfected with a gRNA targeting the noncoding *Rosa26* locus. (v) Cell lines in which dual CRISPR guides were used to generate segmental deletions of genes encoding olfactory receptors on chromosome 1q. (vi) Cell lines in which CRISPR was used to delete a terminal segment on chromosome 1q, eliminating the telomere and decreasing the copy number of 26 out of 968 protein-coding genes on the chromosome.

We performed SMASH-Seq on each control clone that we generated and confirmed that each clone retained an extra copy of chromosome 1q (fig. S7). We then measured the effects of these manipulations on anchorage-independent growth. We found that every control clone maintained the ability to form colonies in soft agar, with some variability between independent clones. Across the three cancer cell lines and 37 different control clones, we observed that the 1q-disomic clones exhibited worse anchorage-independent growth than every control clone that we generated (fig. S8, B to E). These results indicate that the deficiencies in malignant growth exhibited by the 1q-disomic clones are not a result of our use of CRISPR or ganciclovir selection.

Eliminating different cancer aneuploidies produces distinct phenotypic consequences

To further investigate the consequences of inducing aneuploidy loss, we used ReDACT to eliminate the trisomy of either chromosome 7p or 8q from A2058 melanoma cells. SMASH-Seq confirmed the desired aneuploidy-loss events without other karyotypic changes (Fig. 3A, fig. S9, and table S3). As expected, loss of either trisomy 7p or trisomy 8q resulted in a decrease in the expression of genes encoded on the affected chromosomes (Fig. 3B). 7p-disomic and 8q-disomic clones exhibited impaired anchorage-independent growth compared with a panel of control clones, although this defect was not as severe as the defect observed among A2058 1q-disomic clones (Fig. 3C and fig. S10). In vitro doubling times were also closer to wild-type levels for 7p-disomic and 8q-disomic cells compared with 1q-disomic cells (fig. S6D). Finally, we performed subcutaneous injections of the 7p-disomic and the 8q-disomic cells in nude mice, and we found that loss of either the 7p or the 8q trisomy resulted in a moderate decrease in tumor growth (Fig. 3D). At the end of the assay, the wild-type tumors were, on average, two-fold larger than the tumors formed by either 7p-disomic or 8q-disomic cells, compared with a 30-fold difference between A2058 wild-type and 1q-disomic tumors (Fig. 2E). In total, these results indicate that A2058 melanoma cells exhibit a greater degree of "addiction" to the 1q trisomy compared with the trisomies of chromosome 7p or 8q.

To explore the consequences of losing chromosome 8q aneuploidy in a distinct cancer type, we eliminated the 8q trisomy from the colorectal cancer cell line HCT116 (Fig. 3, E and F; fig. S9C; and table S3). Consistent with our observations in A2058, loss of the 8q trisomy decreased but did not fully prevent anchorage-independent growth in HCT116 (Fig. 3G). We then tested xenograft formation in the HCT116 8q-disomic cells, and we observed that one 8q-disomic clone exhibited a moderate defect in tumor growth while a second clone was able to form tumors at levels comparable to the trisomic parental line (Fig. 3H). These results demonstrate that eliminating aneuploid chromosomes has variable effects, depending on the identity of the chromosome and the genetic background of the cancer.

Karyotype evolution and 1q trisomy restoration after aneuploidy loss

A hallmark of oncogene addictions is that loss or inhibition of a driver oncogene results in strong and rapid selection to reestablish oncogenic signaling (*40*). For instance, when epidermal growth factor receptor (EGFR)–driven lung cancers are treated with an EGFR inhibitor, these cancers evolve to acquire specific mutations, including $EGFR^{T790M}$ (which restores EGFR activity) and $KRAS^{G13D}$ (which activates a parallel oncogenic pathway) (*41*). We sought to investigate whether elimination of an "aneuploidy addiction" would also result in evolutionary pressure to restore the lost aneuploidy. We injected 1q-disomic A2058 cells into nude mice and then determined the copy number of chromosome 1q in the resulting xenografts using quantitative polymerase chain reaction (Fig. 4A). We discovered that 65 out of 82 1q-disomic xenografts reacquired an extra copy of chromosome 1q, demonstrating strong selective pressure to regain the initial 1q aneuploidy. We subjected 20 of these post-xenograft clones to SMASH-Seq, and we found that chromosome 1q regain was the only detectable chromosome-scale copy number change (Fig. 4B and fig. S11A). No gross karyotypic changes were observed when the parental 1q-trisomic cells were grown as xenografts (Fig. 4B and fig. S11B). If loss of the chromosome 1q trisomy compromises malignant potential, then we would expect that regaining 1q aneuploidy would restore cell fitness. Consistent with this, we found that cells that had reacquired the 1q trisomy exhibited increased clonogenicity compared with 1q-disomic cells when grown under anchorage-independent conditions (Fig. 4C).

Next, we assessed karyotype evolution after in vivo growth of A2058 7p-disomic and 8q-disomic clones. Interestingly, 17 out of 68 7p-disomic xenografts and 17 out of 63 8q-disomic xenografts were found to exhibit 7p and 8q trisomy regain, respectively (Fig. 4, D and E). These rates of chromosome regain

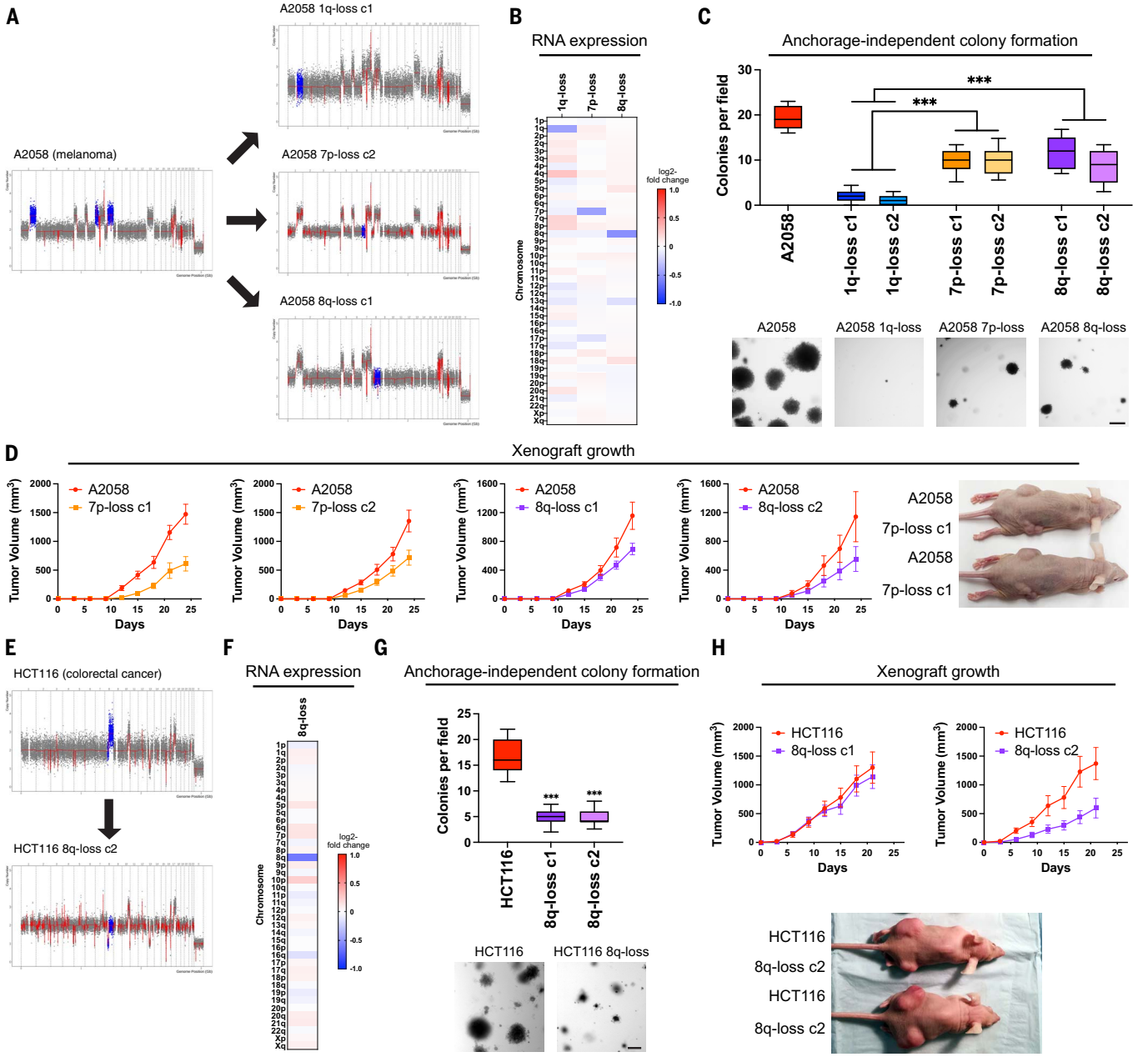


Fig. 3. Variable degrees of addiction to aneuploidy of chromosomes 1q, 7p, and 8q. (**A**) Representative SMASH karyotypes of the 1q-disomic, 7p-disomic, and 8q-disomic clones generated from the melanoma cell line A2058. Trisomy of chromosomes 1q, 7p, and 8q are highlighted in blue in the parental cell line on the left, and the respective targeted chromosome loss is highlighted in blue in the derived clones on the right. A complete list of aneuploidy-loss clones and how they were generated is included in table S3. (**B**) 1q-disomic, 7p-disomic, and 8q-disomic clones in A2058 exhibit decreased RNA expression of genes encoded on the targeted chromosome. RNA expression data were obtained through bulk RNA-seq and represent the average expression of genes by chromosome arm across multiple aneuploidy-loss clones for each targeted chromosome. Data are log2 transformed, normalized to the parental cell line, and adjusted so that the mean expression across all chromosomes is 0. (**C**) 7p-disomic and 8q-disomic clones in A2058 exhibit a milder deficit in anchorage-independent growth than do 1q-disomic clones. The micrographs display representative images of colony formation for the indicated cell lines. (**D**) 7p-disomic and 8q-disomic clones in A2058 exhibit a moderate defect in xenograft growth. Wild-type (7p-trisomic and 8q-trisomic) cells and either 7p-disomic or 8q-disomic cells were injected contralaterally and subcutaneously into immunocompromised mice. The graphs display the mean ± SEM for each trial. Representative mice are shown on the right. (**E**) SMASH karyotype of an 8q-disomic clone generated from the colorectal cancer cell line HCT116. Chromosome 8q is highlighted in blue. (**F**) 8q-disomic clones in HCT116 exhibit decreased RNA expression of genes encoded on chromosome 8q. RNA expression data were obtained through bulk RNA-seq and represent the average expression of genes by chromosome arm across multiple aneuploidy-loss clones for each cell line. Data are log2 transformed, normalized to the parental cell line, and adjusted so that the mean expression across all chromosomes is 0. (**G**) 8q-disomic clones in HCT116 exhibit decreased anchorage-independent growth. The micrographs display representative images of colony formation for the indicated cell lines. (**H**) 8q-disomic clones in HCT116 exhibit variable xenograft growth. 8q-trisomic and 8q-disomic cells were injected contralaterally and subcutaneously into immunocompromised mice. The graphs display the mean ± SEM for each trial. Representative mice are shown below the graphs. For anchorage-independent growth assays in (C) and (G), boxes represent the 25th, 50th, and 75th percentiles of colonies per field, while the whiskers represent the 10th and 90th percentiles. Unpaired *t* test, $n = 15$ fields of view, data from representative trial ($n \geq 2$ total trials). Representative images are shown below. Scale bars, 250 μm. ***$P < 0.0005$.

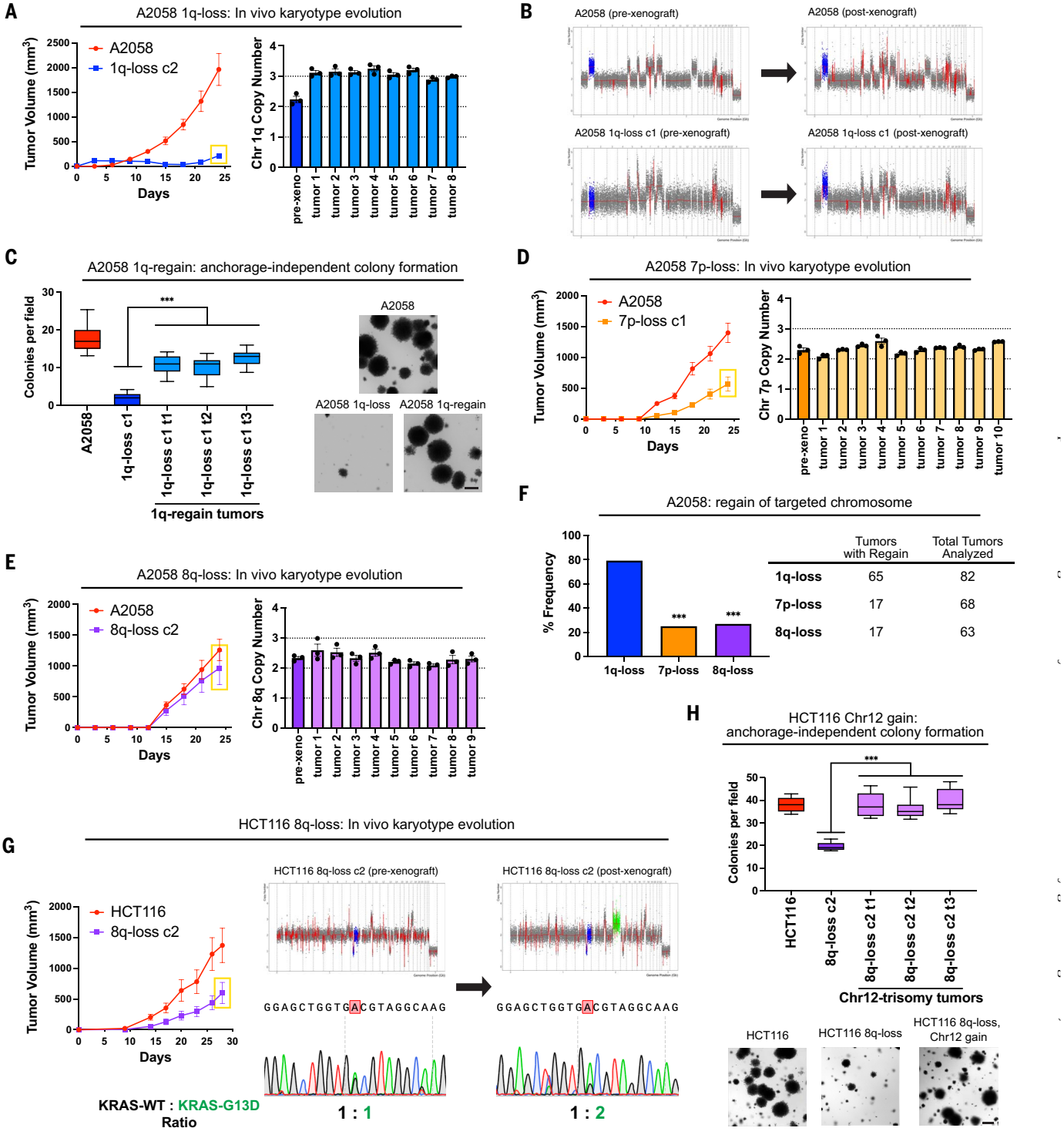


Fig. 4. Cancers rapidly recover chromosome 1q aneuploidy. (**A**) A2058 1q-disomic cells frequently evolve to recover a third copy of chromosome 1q during xenograft growth. (**B**) Representative SMASH karyotypes of A2058 wildtype and 1q-disomic tumors. The initial karyotypes for these lines before the xenograft assay are shown on the left, and karyotypes of tumors after the xenograft assay are shown on the right. Chromosome 1q is highlighted in blue. (**C**) 1q-disomic clones that have evolved to regain 1q trisomy after xenograft growth exhibit increased anchorage-independent growth relative to the pre-xenograft 1q-disomic parental cells. (**D**) Variable evolution of 7p-disomic cells to recover a third copy of chromosome 7p during xenograft growth. (**E**) Variable evolution of 8q-disomic cells to recover a third copy of chromosome 8q during xenograft growth. (**F**) Regain of trisomy 1q occurs more frequently than regain of trisomy 7p or trisomy 8q. Tumors were classified as exhibiting regain if the mean copy number of the targeted chromosome was ≥2.5, as determined through TaqMan copy number assays. n = 213 tumors, chi-squared test. (**G**) HCT116 8q-disomic clones evolve to gain a copy of chromosome 12 during xenograft assays, resulting in the acquisition of an extra copy of the *KRAS*^{G13D} allele. Cell lines were rederived from tumors harvested at the endpoint of xenograft assays and subjected to SMASH karyotyping and Sanger sequencing of *KRAS*. The xenograft growth curve is shown on the left, and representative SMASH karyotype profiles and Sanger sequencing chromatograms

before and after xenograft are shown on the right. Chromosome 8q is highlighted in blue, and chromosome 12 is highlighted in green. (**H**) 8q-disomic clones that have evolved to acquire trisomy of chromosome 12 after xenograft growth exhibit increased anchorage-independent growth relative to the pre-xenograft 8q-disomic parental cells. For copy number profiling in (A), (D), and (E), cell lines were rederived from tumors at the endpoint of the xenograft assays, and chromosome copy number was determined through TaqMan copy number assays. Mean ± SEM, n = 3 probes on targeted chromosome, data from representative trials are shown (n = 2 total trials). The corresponding xenograft assays are shown on the left. For the anchorage-independent growth assays in (C) and (H), the boxes represent the 25th, 50th, and 75th percentiles of colonies per field, while the whiskers represent the 10th and 90th percentiles. Unpaired t test, n = 15 fields of view, data from representative trial ($n \geq 2$ total trials). Representative images are shown on the right. Scale bars, 250 μm. ***$P < 0.0005$.

were significantly lower than the rates that we observed for chromosome 1q ($P < 0.0001$, chi-square test; Fig. 4F). These results suggest that there is moderate selective pressure to restore 7p and 8q trisomies and stronger selective pressure to restore 1q trisomy in A2058.

We then sought to determine whether we could observe evolutionary pressure to restore chromosome 1q aneuploidy when 1q-disomic cells were grown in vitro. Toward that end, we passaged A2058, A2780, and AGS 1q-trisomic and 1q-disomic cancer cell lines for 30 days in culture and then assessed their karyotypes. Similar to our in vivo results, we uncovered multiple instances in which 1q-disomic cells independently regained an extra copy of chromosome 1q over the course of the assay (fig. S12). We reasoned that there were two possible sources for the regained chromosome: The third copy of 1q could result from the missegregation of an endogenous copy of chromosome 1q, or the disomic cell line could have been outcompeted by an exogenous population of trisomic cells (for instance, from metastatic colonization when 1q-disomic and 1q-trisomic xenografts were grown in the same mouse). To differentiate between these possibilities, we identified single-nucleotide polymorphisms on 1q that were heterozygous in the parental cell line and that exhibited loss of heterozygosity after 1q elimination. We reasoned that if loss of heterozygosity was maintained after 1q regain, then that would be evidence for an endogenous missegregation event, whereas the reappearance of heterozygosity would be evidence of an exogenous cell population (fig. S13A). Sanger sequencing analysis revealed that both potential causes of 1q regain were observed during these evolution experiments (fig. S13, B and C). In both cases, 1q regain correlated with restored anchorage-independent growth relative to the 1q-disomic clones, further verifying the link between 1q copy number and cell fitness (fig. S13, D and E).

Finally, we assessed karyotype evolution in xenografts produced by 8q-disomic HCT116 cells (Fig. 4G). We found that 0 out of 13 tumors regained the trisomy of chromosome 8q, but 7 out of 13 tumors gained a de novo trisomy of chromosome 12. HCT116 cells are driven by a heterozygous $KRAS^{G13D}$ mutation (*16*), and *KRAS* is encoded on chromosome 12. Sanger sequencing analysis revealed that every chromosome 12–trisomic tumor had amplified the copy of chromosome 12 harboring the mutant $KRAS^{G13D}$ allele (Fig. 4G). Increasing dosage of mutant *KRAS* has previously been associated with enhanced tumor fitness (*42*), and we observed that these chromosome 8q–disomic and chromosome 12–trisomic cells exhibited superior anchorage-independent growth relative to the chromosome 8q–disomic and chromosome 12–disomic pre-xenograft population (Fig. 4H). In total, these results suggest that aneuploidy loss creates strong selective pressure for karyotype evolution, and the effects of aneuploidy loss can be suppressed in cis (by regaining the lost chromosome) or in trans (by acquiring a beneficial secondary alteration).

Chromosome 1q aneuploidy suppresses p53 signaling by increasing MDM4 *expression*

We sought to uncover the biological mechanism underlying the addiction to chromosome 1q aneuploidy. RNA sequencing (RNA-seq) analysis revealed that elimination of the 1q trisomy caused up-regulation of tumor suppressor p53 target genes in A2780 and MCF10A, which are both wild type for *TP53* (Fig. 5, A and B, and fig. S14, A and B). Western blotting confirmed that 1q-disomic clones exhibited increased phosphorylation of p53 at serine-15 and increased expression of the canonical p53 target p21 (Fig. 5C) (*43*). These results were not a by-product of CRISPR mutagenesis, as A2780 cells harboring a CRISPR-mediated integration of *HSV-TK* did not display evidence of p53 activation (Fig. 5, B and C). Additionally, 1q-disomic A2780 and MCF10A cells exhibited a delay in the G1 phase of the cell cycle and increased senescence-associated β-galactosidase staining, both of which are associated with p53-mediated tumor suppression (fig. S14, C to F) (*44*). These results suggest that the chromosome 1q trisomy inhibits p53 signaling and that elimination of this trisomy antagonizes malignant growth at least in part by triggering p53 activation.

To explore whether p53 inhibition is a common consequence of chromosome 1q gains, we examined our prior analysis of aneuploidy-mutation mutual exclusivity in cancer genomes (table S1). Out of 14,383 aneuploidy–gene mutation pairs, the single strongest instance of mutual exclusivity was between 1q gains and *TP53* mutations (Fig. 5D). Next, we applied a classification algorithm capable of predicting cancers that lack p53 function on the basis of their transcriptional profiles (*45*). As expected, cancers from TCGA with nonsynonymous *TP53* mutations scored higher with this classifier than cancers with wild-type *TP53* (Fig. 5E). Considering only tumors with wild-type *TP53*, we calculated the association between the p53 status classifier and every possible chromosome arm gain in TCGA. Across all chromosomes, 1q gains exhibited the strongest correlation with the p53-loss signature (Fig. 5, F and G). Among tumors with wild-type *TP53*, gains of chromosome 1q were associated with lower expression of the p53 target genes *CDKN1A* (p21), *GADD45A*, and *RRM2B* (Fig. 5H) (*43*). In total, these results indicate that gaining chromosome 1q phenocopies the effects of p53 mutations and suppresses p53 activity in human tumors.

We sought to discover the gene(s) on chromosome 1q responsible for inhibiting p53 signaling. We noted that *MDM4*, a negative regulator of p53 activity, is located on 1q32 (*46*). *MDM4* expression increased with chromosome 1q copy number and higher *MDM4* expression correlated with the p53-loss transcriptional signature (fig. S15, A and B). To uncover whether *MDM4* is directly responsible for the 1q-aneuploidy addiction observed in A2780, we first used CRISPR interference (CRISPRi) to down-regulate *MDM4* expression without fully ablating it (*47*). In A2780 competition assays, we observed that down-regulating *MDM4* impaired cell fitness relative to A2780 cells in which *AAVS1* or *PIP5K1A*, an unrelated gene on chromosome 1q, were down-regulated (Fig. 5I) (*48*). Next, we used a two-guide strategy to delete a single copy of *MDM4* in an otherwise trisomic background (Fig. 5, J and K, and fig. S15, C and D). We found that the subsequent A2780 $MDM4^{+/+/KO}$ clones down-regulated *MDM4* and up-regulated p53 target genes, comparable to the effects observed in cells lacking the entire 1q trisomy (Fig. 5L and fig. S15E). We then tested the colony formation ability of $MDM4^{+/+/KO}$ clones, and we discovered that losing a single copy of *MDM4* decreased anchorage-independent growth (Fig. 5M). Subsequently, we performed the converse experiment: We cloned *MDM4* cDNA under the control of a doxycycline-inducible promoter and transduced the construct into both 1q-trisomic and 1q-disomic cells. We found that moderate (1.7-fold) overexpression of *MDM4* was sufficient to cause an increase in anchorage-independent growth in

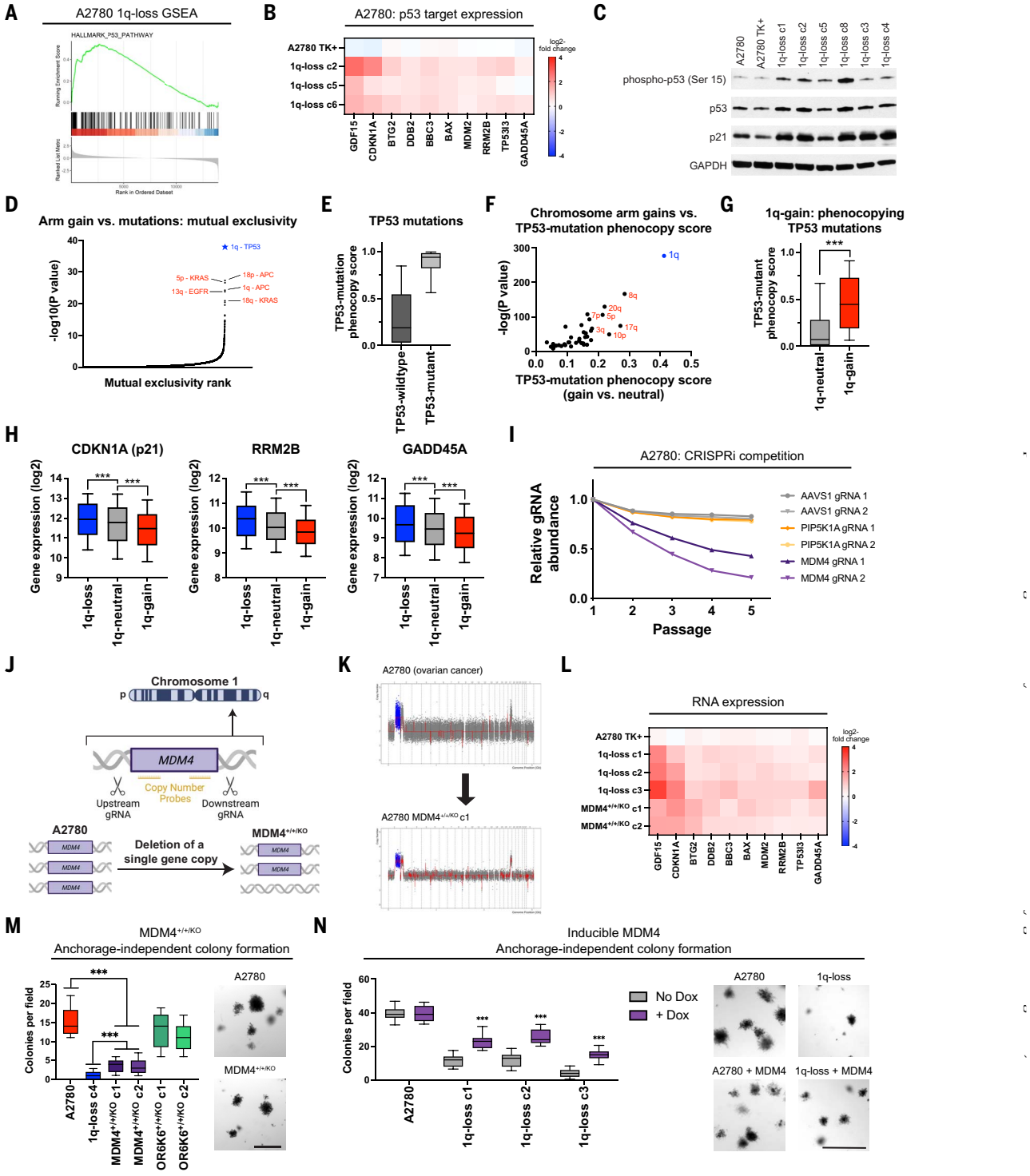


Fig. 5. A single extra copy of *MDM4* suppresses p53 signaling and contributes to the 1q trisomy addiction. (**A**) GSEA analysis of A2780 RNA-seq data reveals up-regulation of the p53 pathway in the 1q-disomic clones, relative to the parental trisomy. (**B**) A heatmap displaying the up-regulation of 10 p53 target genes in A2780 1q-disomic clones. The TK+ clone indicates a clone that harbors the CRISPR-mediated integration of the *HSV-TK* transgene but that was not treated to induce chromosome 1q-loss. (**C**) Western blot analysis demonstrating activation of p53 signaling in 1q-disomic clones. Glyceraldehyde phosphate dehydrogenase (GAPDH) was analyzed as a loading control. The TK+ clone indicates a clone that harbors the CRISPR-mediated integration of the *HSV-TK* transgene but that was not treated to induce chromosome 1q-loss. (**D**) A waterfall plot highlighting the most-significant instances of mutual exclusivity between chromosome arm gains and mutations in cancer-associated genes. The complete dataset for mutual exclusivity and co-occurrence is included in table S1. (**E**) Boxplots displaying the *TP53*-mutation phenocopy signature (*45*) in cancers from TCGA, split according to whether the cancers harbor a nonsynonymous mutation in *TP53*. (**F**) A scatterplot comparing the association between chromosome arm gains and the *TP53*-mutation phenocopy signature (*45*) in *TP53* wild-type cancers from TCGA. Cancers with chromosome 1q gains are highlighted in blue. (**G**) Boxplots displaying the *TP53*-mutation phenocopy signature (*45*) in cancers from TCGA, split according to whether

tumors harbor a gain of chromosome 1q. Only *TP53* wild-type cancers are included in this analysis. (**H**) Boxplots displaying the expression of three p53 target genes—*CDKN1A* (p21), *RRM2B*, and *GADD45A*—in cancers from TCGA split according to the copy number of chromosome 1q. Only *TP53* wild-type cancers are included in this analysis. (**I**) A CRISPRi competition assay demonstrates that gRNAs targeting *MDM4* drop out over time in A2780 cells. In contrast, gRNAs targeting *AAVS1* and *PIP5K1A*, another gene encoded on chromosome 1q, exhibit minimal depletion. (**J**) A schematic displaying the strategy for using paired CRISPR gRNAs to delete a single copy of *MDM4* in a cell line with a trisomy of chromosome 1q. (**K**) SMASH karyotype demonstrating maintenance of the chromosome 1q trisomy in an $MDM4^{+/+/KO}$ clone. Chromosome 1q is highlighted in blue. (**L**) 1q-disomic clones and $MDM4^{+/+/KO}$ clones in A2780 exhibit comparable up-regulation of p53 transcriptional targets, as determined through TaqMan gene expression assays. (**M**) $MDM4^{+/+/KO}$ clones exhibit decreased anchorage-independent growth relative to the $MDM4^{+/+/+}$ parental cell line. (**N**) Induction of *MDM4* cDNA in 1q-disomic clones in A2780 increases anchorage-independent growth. For the graphs in (E), (G), (H), (M), and (N), the boxplots represent the 25th, 50th, and 75th percentiles of the indicated data, while the whiskers represent the 10th and 90th percentiles of the indicated data. For the soft agar experiments in (M) and (N), the data are from $n = 15$ fields of view, and a representative trial is shown ($n \geq 2$ total trials). Scale bars, 250 μm. ***$P < 0.0005$.

the 1q-disomic cells, whereas this same treatment did not affect the 1q-trisomic cells (Fig. 5N and fig. S15F).

Finally, to investigate the role of p53 as a mediator of 1q aneuploidy addiction from an orthogonal approach, we used CRISPR to delete the *TP53* gene in A2780 1q-disomic and 1q-trisomic cells (fig. S16, A and B). We discovered that loss of *TP53* rescued the G1 delay and enhanced anchorage-independent growth in 1q-disomic cells (fig. S16, C and D). The magnitude of the increase in colony formation was significantly greater in the 1q-disomic cells than in the 1q-trisomic cells (4-fold versus 1.5-fold; $P < 0.0001$, *t* test) (fig. S16D). In total, these results indicate that chromosome 1q gains are a mechanism by which *TP53* wild-type cancers can suppress p53 activity, and this suppression occurs because of the increased expression of *MDM4*.

BCL9 promotes the growth of 1q-aneuploid cancers through a p53-independent mechanism

We noted that the deletion of *TP53* and the overexpression of *MDM4* in 1q-disomic clones did not fully restore anchorage-independent growth to 1q-trisomic levels (Fig. 5N and fig. S16D). We therefore hypothesized that there were additional dosage-sensitive genes encoded on chromosome 1q that promote the fitness of 1q-aneuploid cancers. To uncover these genes, we assembled a panel of 16 1q genes that have previously been associated with tumorigenesis, including *MDM4* as a positive control, and we conducted CRISPRi competition assays to assess the effects of down-regulating each gene in 1q-trisomic A2780 cells (fig. S17A). gRNAs targeting six genes, including *MDM4*, exhibited a mean depletion of >1.75-fold in cellular competitions (fig. S17B). We cloned cDNAs for these six genes into doxycycline-inducible vectors and transduced them into A2780 1q-disomic cells. We found that overexpression of three of these genes increased anchorage-independent growth: *MDM4*, the antiapoptotic gene *MCL1*, and the Wnt/β-catenin signaling gene *BCL9* (fig. S17C). We then conducted two further analyses to test whether the effects of the latter two genes were independent of the MDM4/p53 pathway. First, we cotransduced 1q-disomic cells with vectors to overexpress both *MDM4* and *BCL9* or *MDM4* and *MCL1*, and next we assessed the effects of overexpressing *BCL9* or *MCL1* in *TP53-KO* 1q-disomic cells. We found that *MCL1* had no effect on anchorage-independent growth in 1q-disomic cells that lacked *TP53*, and coexpressing *MCL1* and *MDM4* in p53 wild-type 1q-disomic cells did not increase clonogenicity beyond the effects of expressing *MCL1* alone (fig. S17, D and E). As both *MCL1* and *TP53* control apoptosis, we speculate that the overexpression of *MCL1* and *MDM4* are to some extent epistatic with one another (*49*).

In contrast, *BCL9* expression enhanced anchorage-independent growth in *TP53-KO* cells, and *BCL9-MDM4* coexpression resulted in more colony formation compared with *BCL9* alone (fig. S17, D and E). *BCL9* encodes an adaptor protein that binds to nuclear β-catenin and enhances β-catenin–mediated transcriptional activity (*50*). We found that eliminating the trisomy of chromosome 1q reduced the expression of *BCL9* as well as *AXIN2* and *LGR5*, which are canonical targets of the Wnt/β-catenin pathway (*51*) (fig. S18A). Gene set enrichment analysis (GSEA) revealed a general decrease in the expression of transcripts associated with the Wnt/β-catenin pathway in 1q-disomic versus 1q-trisomic cells (fig. S18B). Ectopic overexpression of *BCL9* increased the expression of *AXIN2* and *LGR5* (fig. S18C). In human cancers, chromosome 1q gains were associated with higher levels of *BCL9* expression, and high *BCL9* was associated with the up-regulation of *AXIN2* and *LGR5* (fig. S18, D and E). Mutations in *CTNNB1*, which encodes β-catenin, were overrepresented in cancers with 1q gains (fig. S18F). Finally, 1q aneuploidy was associated with higher β-catenin activity in several cancer types, including ovarian cancer, hepatocellular carcinoma, and rectal adenocarcinoma (fig. S18G). In total, these results indicate that chromosome 1q aneuploidy can enhance oncogenic Wnt/β-catenin signaling through the up-regulation of *BCL9* and that increased *BCL9* expression promotes cancer cell fitness in a p53-independent manner.

Aneuploidy addictions create collateral therapeutic vulnerabilities

The oncogene addiction hypothesis is the conceptual foundation for the use of targeted therapies in cancer (*40*). Correspondingly, we sought to uncover whether aneuploidy addictions could also represent a therapeutic vulnerability for certain cancers. We noted that chromosome 1q harbors the *UCK2* gene, which encodes a pyrimidine salvage kinase that controls a rate-limiting step in the activation of certain toxic nucleotide analogs, including RX-3117 and 3-deazauridine (Fig. 6A) (*52*, *53*). We found that *UCK2* is overexpressed in human cancers that contain extra copies of chromosome 1q, and elimination of the chromosome 1q trisomy consistently decreased UCK2 protein expression in our engineered cell lines (Fig. 6, B and C). We therefore investigated whether gaining chromosome 1q could create a collateral sensitivity to UCK2-dependent nucleotide analogs.

First, as the mechanism of many cancer drugs is poorly understood (*54*, *55*), we sought to verify that the cytotoxicity of RX-3117 and 3-deazauridine requires *UCK2* expression. We used CRISPR to delete *UCK2* in the haploid HAP1 cell line, and we confirmed that *UCK2*-knockout cells were highly resistant to both compounds (fig. S19, A to C). Next, we tested the effects of RX-3117 and 3-deazauridine in our engineered 1q-trisomic and 1q-disomic cell lines. We found that A2780 and MCF10A cells harboring a trisomy of chromosome 1q were more sensitive to both compounds compared with isogenic cells containing two copies of chromosome 1q (Fig. 6D). This effect was specific for UCK2 substrates, as the 1q-trisomic cells did not exhibit greater sensitivity to UCK2-independent nucleotide poisons and other cancer drugs (fig. S19D). Furthermore, deletion of a single copy of *UCK2* in 1q-trisomic A2780 cells was sufficient to decrease sensitivity to RX-3117, whereas ectopic overexpression of *UCK2* cDNA in 1q-disomic A2780 cells was sufficient to increase sensitivity to RX-3117 (fig. S19, E to H). However, we did not detect any difference in RX-3117 sensitivity between 1q-trisomic and 1q-disomic A2058 and AGS cells (fig. S20, A and B). As A2780 and MCF10A

harbor wild-type *TP53*, whereas AGS and A2058 exhibit compromised p53 signaling, we speculate that *TP53* status or the expression of other related proteins may also influence the response to UCK2 substrates.

To determine whether 1q copy number changes generated without using ReDACT could also increase sensitivity to UCK2 substrates, we transiently treated near-diploid DLD1 colon cancer cells with an inhibitor of the spindle checkpoint kinase Mps1, and we isolated a clone that harbored a trisomy of chromosome 1 (fig. S20, C and D). DLD1 trisomy-1 cells were significantly more sensitive to RX-3117 compared with the parental DLD1 cells ($P < 0.005$, *t* test; fig. S20E). Finally, RX-3117 and 3-deazauridine have been screened across the NCI-60 cell line panel, and we found that higher *UCK2* expression correlates with greater sensitivity to both compounds (fig. S20F) (*56*). In total, these results indicate that 1q gains induce a collateral sensitivity to certain nucleotide analogs by increasing the expression of *UCK2*.

We hypothesized that we could use the greater sensitivity of 1q-trisomic cells to UCK2 substrates to redirect cellular evolution away from aneuploidy and toward a disomic state with lower malignant potential (Fig. 6E). We mixed fluorescently labeled 1q-trisomic and 1q-disomic MCF10A cells at a ratio of 20:80 and then co-cultured the two cell populations. After 9 days of growth in drug-free medium,

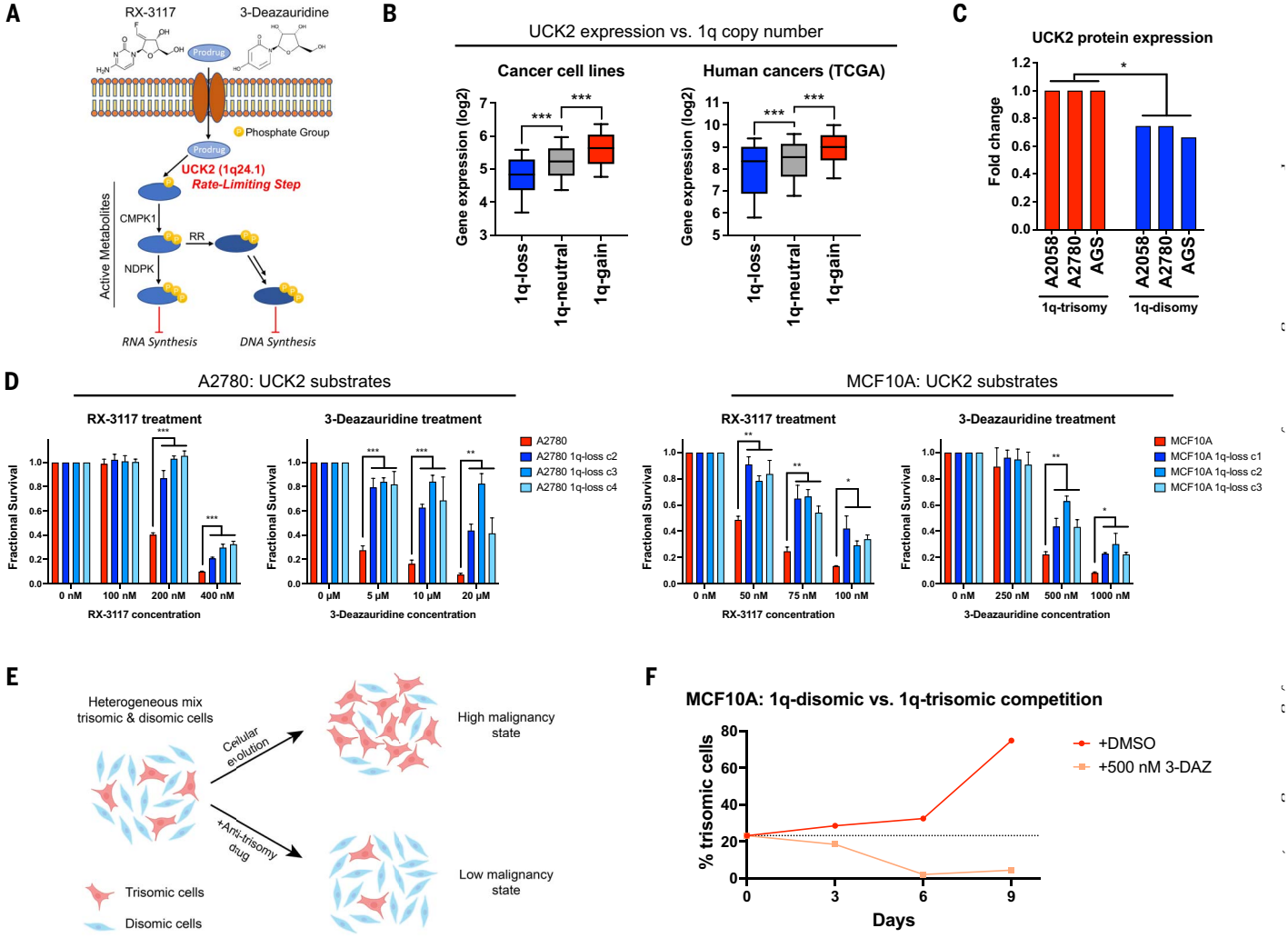


Fig. 6. Gaining chromosome 1q increases sensitivity to UCK2 substrates. (**A**) A schematic of the metabolism of two pyrimidine analogs, RX-3117 and 3-deazauridine. UCK2, a kinase encoded on chromosome 1q, phosphorylates these compounds to produce cytotoxic derivatives that can poison DNA and RNA synthesis. (**B**) Boxplots displaying the expression of *UCK2* from the Cancer Cell Line Encyclopedia (CCLE, left) and TCGA (right), divided according to the copy number of chromosome 1q. The boxplots represent the 25th, 50th, and 75th percentiles of the indicated data, while the whiskers represent the 10th and 90th percentiles of the indicated data. Data were analyzed using unpaired *t* tests; $n = 10{,}331$ samples from TCGA and 942 samples from CCLE. (**C**) Expression of UCK2 protein in cancer cell lines with 1q trisomies or after aneuploidy elimination. (**D**) Cellular sensitivity of A2780 and MCF10A treated with different concentrations of RX-3117 or 3-deazauridine. Mean ± SEM, data from representative trials are shown ($n \geq 3$ total trials). (**E**) A schematic displaying cellular competition between trisomic and disomic cells. Under normal conditions, certain trisomies enhance cellular fitness, allowing these cells to overtake the population and enhance malignant growth (top). However, treatment with an "anti-trisomy" compound could selectively impair the growth of the aneuploid cells, keeping the population in a low-malignant state (bottom). (**F**) A cellular competition between fluorescently labeled MCF10A 1q-trisomic and unlabeled 1q-disomic cells. These cells were mixed at a ratio of 20% to 80% and then cultured in either dimethyl sulfoxide (DMSO) or 3-deazauridine. While the trisomic cells quickly dominate the population in drug-free medium, treatment with 3-deazauridine prevents the outgrowth of the 1q-trisomic subpopulation. Data from representative trials are shown ($n = 2$ total trials). *$P < 0.05$, **$P < 0.005$, ***$P < 0.0005$.

the trisomic cells had expanded to make up 75% of the culture, but when the same cell populations were grown in the presence of 500 nM 3-deazauridine, the trisomic population decreased to just 4% of the final culture (Fig. 6F). We conclude that trisomy-selective compounds can be used to manipulate cellular evolution to prevent the outgrowth of malignant aneuploid cells in a premalignant setting.

Finally, we investigated whether we could generalize this approach to identify compounds that exhibit selective toxicity against other aneuploidies. To nominate drugs that could be used to target chromosome 7p, we assessed the PRISM dataset of 4518 compounds tested against 578 cancer cell lines (*57*). We calculated the correlation between the expression of every gene encoded on chromosome 7p and the sensitivity to each drug (fig. S21A). One of the strongest relationships that we found was between expression of the gene *AHR*, which encodes a ligand-activated transcription factor, and sensitivity to the drug CGS-15943 (fig. S21B). It has previously been reported that CGS-15943 binds to AHR and causes it to up-regulate the expression of proapoptotic genes, providing a potential mechanistic explanation for this result (fig. S21C) (*58*). *AHR* expression was up-regulated in human cancer cell lines and tumors that are trisomic for chromosome 7p, whereas eliminating the 7p trisomy in our A2058 cell line model decreased the expression of *AHR* (fig. S21D). Consistent with our hypothesis, we found that A2058 cells with a 7p trisomy were moderately but significantly more sensitive to CGS-15943 compared with 7p-disomic A2058 cells ($P < 0.0005$, t test; fig. S21E). We conclude that by using a combination of computational and experimental approaches we can uncover compounds that exhibit greater activity toward cancer cells with specific aneuploidies.

Discussion

In this work, we eliminated endogenous aneuploidies from established cancer cell lines and revealed that the removal of trisomic chromosomes compromises cancer-like growth. We posit that these phenotypes are due specifically to the loss of the aneuploid chromosome and are not a by-product of CRISPR selection or the elimination of point mutations encoded on the targeted chromosome (discussed in more detail in the supplementary text) (*59*). Due to the similarity between our observations and the previously described oncogene addiction phenomenon (*17*), we suggest that in certain circumstances cancers may also be "addicted" to the aneuploidy found in their genomes. We speculate that during tumor evolution, certain aneuploidies can provide context-specific benefits that enhance tumorigenesis. For instance, we showed that chromosome 1q gains are an early event during cancer development, and we demonstrated that *MDM4* and *BCL9* are dosage-sensitive genes on 1q that enhance malignant growth. In cells that already harbor *TP53* mutations or in cancer types that are not driven by WNT signaling, the beneficial effects of gaining chromosome 1q may be outweighed by the detrimental effects of overexpressing hundreds of other 1q genes.

MDM4 and many other genes have tumor-promoting properties when highly overexpressed (*60*, *61*). For instance, *MDM4* is focally amplified in ~1% of cancers in TCGA, and strong overexpression of *MDM4* via retrovirus immortalizes primary cells and renders them sensitive to Ras-mediated transformation (*62*). In this study, we demonstrated that a single extra copy of *MDM4* is sufficient to suppress the expression of p53 target genes and promote oncogenic growth. Our results are consistent with a recent report showing that low-level overexpression of *MDM4* can enhance fitness in hematopoietic cell competitions (*63*). The overlap between single-copy dosage-sensitive genes such as *MDM4* and genes found to have tumor-promoting properties when highly overexpressed is at present unknown. Furthermore, our work demonstrates that *MDM4* overexpression alone is insufficient to fully recapitulate the oncogenic effects of 1q aneuploidy, and we established that *BCL9* is a second dosage-sensitive 1q gene that enhances fitness in a *TP53*-independent manner. As arm-length aneuploidy is more common in tumor genomes than focal gene amplifications (*19*), we expect that recurrently gained chromosomes harbor multiple independent loci such as *MDM4* and *BCL9* that cooperate to drive tumor development.

Finally, our results raise the exciting possibility that "aneuploidy addictions" may represent a therapeutic vulnerability in cancer. Previous attempts to target aneuploidy have focused on phenotypes that are shared across highly-aneuploid cells, such as alterations in spindle geometry (*64*, *65*). Here, we sought to develop an approach to take advantage of the genes that are encoded on an aneuploid chromosome, thereby allowing chromosome-specific targeting. In particular, we hypothesized that the overexpression of specific genes—for instance, drug-importer pumps or enzymes required for a prodrug's activation—could sensitize cancers to compounds that are otherwise better tolerated in euploid tissue. We demonstrated that gaining chromosome 1q creates a collateral vulnerability to the nucleotide analogs RX-3117 and 3-deazauridine owing to the overexpression of the kinase UCK2. Notably, RX-3117 has been tested in phase 2A clinical trials, but without the use of any genomic biomarkers for patient selection (*66*). High *UCK2* expression has been proposed as a potential sensitivity biomarker for RX-3117, and we speculate that patients whose tumors harbor gains of chromosome 1q may exhibit strong responses because of the constitutive overexpression of *UCK2* (*67*). More broadly, compounds whose anticancer function is enhanced by genes encoded on aneuploid chromosomes could be used to direct cellular evolution away from certain aneuploidies and toward the lower-malignancy diploid state.

Materials and methods summary

The identities of all cell lines used in this study were confirmed using short tandem repeat profiling. CRISPR plasmids were cloned as described in (*68*). Chromosome copy number analysis was performed as described in (*36*). Complete methods are provided in the supplementary materials (*59*).

REFERENCES AND NOTES

1. A. M. Taylor *et al*., Genomic and functional approaches to understanding cancer aneuploidy. *Cancer Cell* **33**, 676–689.e3 (2018). doi: 10.1016/j.ccell.2018.03.007; pmid: 29622463
2. D. J. Gordon, B. Resio, D. Pellman, Causes and consequences of aneuploidy in cancer. *Nat. Rev. Genet.* **13**, 189–203 (2012). doi: 10.1038/nrg3123; pmid: 22269907
3. T. Boveri, Concerning the origin of malignant tumours by Theodor Boveri. Translated and annotated by Henry Harris. *J. Cell Sci.* **121** (suppl. 1), 1–84 (2008). doi: 10.1242/jcs.025742; pmid: 18089652
4. P. A. Hardy, H. Zacharias, Reappraisal of the Hansemann-Boveri hypothesis on the origin of tumors. *Cell Biol. Int.* **29**, 983–992 (2005). doi: 10.1016/j.cellbi.2005.10.001; pmid: 16314117
5. A. Vasudevan *et al*., Aneuploidy as a promoter and suppressor of malignant growth. *Nat. Rev. Cancer* **21**, 89–103 (2021). doi: 10.1038/s41568-020-00321-1; pmid: 33432169
6. J. M. Sheltzer, A. Amon, The aneuploidy paradox: Costs and benefits of an incorrect karyotype. *Trends Genet.* **27**, 446–453 (2011). doi: 10.1016/j.tig.2011.07.003; pmid: 21872963
7. J. M. Sheltzer *et al*., Single-chromosome gains commonly function as tumor suppressors. *Cancer Cell* **31**, 240–255 (2017). doi: 10.1016/j.ccell.2016.12.004; pmid: 28089890
8. B. A. A. Weaver, A. D. Silk, C. Montagna, P. Verdier-Pinard, D. W. Cleveland, Aneuploidy acts both oncogenically and as a tumor suppressor. *Cancer Cell* **11**, 25–36 (2007). doi: 10.1016/j.ccr.2006.12.003; pmid: 17189716
9. T. Davoli *et al*., Cumulative haploinsufficiency and triplosensitivity drive aneuploidy patterns and shape the cancer genome. *Cell* **155**, 948–962 (2013). doi: 10.1016/j.cell.2013.10.011; pmid: 24183448
10. L. M. Sack *et al*., Profound tissue specificity in proliferation control underlies cancer drivers and aneuploidy patterns. *Cell* **173**, 499–514.e23 (2018). doi: 10.1016/j.cell.2018.02.037; pmid: 29576454
11. D. Zimonjic, M. W. Brooks, N. Popescu, R. A. Weinberg, W. C. Hahn, Derivation of human tumor cells in vitro without widespread genomic instability. *Cancer Res.* **61**, 8838–8844 (2001). pmid: 11751406
12. H. Hasle, J. M. Friedman, J. H. Olsen, S. A. Rasmussen, Low risk of solid tumors in persons with Down syndrome. *Genet. Med.* **18**, 1151–1157 (2016). doi: 10.1038/gim.2016.23; pmid: 27031084
13. K. Bister, Discovery of oncogenes: The advent of molecular cancer research. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 15259–15260 (2015). doi: 10.1073/pnas.1521145112; pmid: 26644573
14. T. A. O'Loughlin, L. A. Gilbert, Functional genomics for cancer research: Applications in vivo and in vitro. *Annu. Rev. Cancer Biol.* **3**, 345–363 (2019). doi: 10.1146/annurev-cancerbio-030518-055742
15. F. Bunz *et al*., Requirement for p53 and p21 to sustain G_2 arrest after DNA damage. *Science* **282**, 1497–1501 (1998). doi: 10.1126/science.282.5393.1497; pmid: 9822382
16. S. Shirasawa, M. Furuse, N. Yokoyama, T. Sasazuki, Altered growth of human colon cancer cell lines disrupted at activated Ki-ras. *Science* **260**, 85–88 (1993). doi: 10.1126/science.8465203; pmid: 8465203

17. I. B. Weinstein, Addiction to oncogenes—the Achilles heal of cancer. *Science* **297**, 63–64 (2002). doi: 10.1126/science.1073096; pmid: 12098689
18. A. M. Waters, C. J. Der, KRAS: The critical driver and therapeutic target for pancreatic cancer. *Cold Spring Harb. Perspect. Med.* **8**, a031435 (2018). doi: 10.1101/cshperspect.a031435; pmid: 29229669
19. T. I. Zack *et al*., Pan-cancer patterns of somatic copy number alteration. *Nat. Genet.* **45**, 1134–1140 (2013). doi: 10.1038/ng.2760; pmid: 24071852
20. Z. Wang *et al*., Evolving copy number gains promote tumor expansion and bolster mutational diversification. bioRxiv 2022.06.14.495959 [Preprint] (2022). https://doi.org/10.1101/2022.06.14.495959.
21. L. R. Yates *et al*., Genomic evolution of breast cancer metastasis and relapse. *Cancer Cell* **32**, 169–184.e7 (2017). doi: 10.1016/j.ccell.2017.07.005; pmid: 28810143
22. N. K. Hayward *et al*., Whole-genome landscapes of major melanoma subtypes. *Nature* **545**, 175–180 (2017). doi: 10.1038/nature22071; pmid: 28467829
23. C. Paterson, H. Clevers, I. Bozic, Mathematical model of colorectal cancer initiation. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 20681–20688 (2020). doi: 10.1073/pnas.2003771117; pmid: 32788368
24. N. McGranahan *et al*., Clonal status of actionable driver events and the timing of mutational processes in cancer evolution. *Sci. Transl. Med.* **7**, 283ra54 (2015). doi: 10.1126/scitranslmed.aaa1408; pmid: 25877892
25. J. Cisowski, M. O. Bergo, What makes oncogenes mutually exclusive? *Small GTPases* **8**, 187–192 (2017). doi: 10.1080/21541248.2016.1212689; pmid: 27416373
26. B. Nguyen *et al*., Genomic characterization of metastatic patterns from prospective clinical sequencing of 25,000 patients. *Cell* **185**, 563–575.e11 (2022). doi: 10.1016/j.cell.2022.01.003; pmid: 35120664
27. R. A. Hagenson, sheltzer-lab/aneuploidy-addictions: Code cleanup, version 1.0.2, Zenodo (2023); https://doi.org/10.5281/zenodo.7927122.
28. K. H. Stopsack *et al*., Aneuploidy drives lethal progression in prostate cancer. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 11390–11395 (2019). doi: 10.1073/pnas.1902645116; pmid: 31085648
29. H. Hieronymus *et al*., Tumor copy number alteration burden is a pan-cancer prognostic factor associated with recurrence and death. *eLife* **7**, e37294 (2018). doi: 10.7554/eLife.37294; pmid: 30178746
30. D. A. Lukow *et al*., Chromosomal instability accelerates the evolution of resistance to anti-cancer therapies. *Dev. Cell* **56**, 2427–2439.e4 (2021). doi: 10.1016/j.devcel.2021.07.009; pmid: 34352222
31. J. Smith, joan-smith/comprehensive-tcga-survival: Genome-wide identification and analysis of prognostic features in human cancers, version 1.0.0, Zenodo (2022); https://doi.org/10.5281/zenodo.6632067.
32. C. Fillat, M. Carrió, A. Cascante, B. Sangro, Suicide gene therapy mediated by the Herpes Simplex virus thymidine kinase gene/Ganciclovir system: Fifteen years of application. *Curr. Gene Ther.* **3**, 13–26 (2003). doi: 10.2174/1566523033347426; pmid: 12553532
33. N. Uno *et al*., CRISPR/Cas9-induced transgene insertion and telomere-associated truncation of a single human chromosome for chromosome engineering in CHO and A9 cells. *Sci. Rep.* **7**, 12739 (2017). doi: 10.1038/s41598-017-10418-7; pmid: 28986519
34. M. V. Zuccaro *et al*., Allele-specific chromosome removal after Cas9 cleavage in human embryos. *Cell* **183**, 1650–1664.e15 (2020). doi: 10.1016/j.cell.2020.10.025; pmid: 33125898
35. S. Papathanasiou *et al*., Whole chromosome loss and genomic instability in mouse embryos after CRISPR-Cas9 genome editing. *Nat. Commun.* **12**, 5855 (2021). doi: 10.1038/s41467-021-26097-y; pmid: 34615869
36. Z. Wang *et al*., SMASH, a fragmentation and sequencing method for genomic copy number analysis. *Genome Res.* **26**, 844–851 (2016). doi: 10.1101/gr.201491.115; pmid: 27197213
37. K. M. Schukken, J. M. Sheltzer, Extensive protein dosage compensation in aneuploid human cancers. *Genome Res.* **32**, 1254–1270 (2022). doi: 10.1101/gr.276378.121; pmid: 35701073
38. S. I. Shin, V. H. Freedman, R. Risser, R. Pollack, Tumorigenicity of virus-transformed cells in nude mice is correlated specifically with anchorage independent growth in vitro. *Proc. Natl. Acad. Sci. U.S.A.* **72**, 4435–4439 (1975). doi: 10.1073/pnas.72.11.4435; pmid: 172908
39. L. Chin *et al*., Essential role for oncogenic Ras in tumour maintenance. *Nature* **400**, 468–472 (1999). doi: 10.1038/22788; pmid: 10440378
40. D. Torti, L. Trusolino, Oncogene addiction as a foundational rationale for targeted anti-cancer therapy: Promises and perils. *EMBO Mol. Med.* **3**, 623–636 (2011). doi: 10.1002/emmm.201100176; pmid: 21953712
41. D. Westover, J. Zugazagoitia, B. C. Cho, C. M. Lovly, L. Paz-Ares, Mechanisms of acquired resistance to first- and second-generation EGFR tyrosine kinase inhibitors. *Ann. Oncol.* **29** (suppl. 1), i10–i19 (2018). doi: 10.1093/annonc/mdx703; pmid: 29462254
42. M. R. Burgess *et al*., KRAS allelic imbalance enhances fitness and modulates MAP kinase dependence in cancer. *Cell* **168**, 817–829.e15 (2017). doi: 10.1016/j.cell.2017.01.020; pmid: 28215705
43. M. Fischer, Census and evaluation of p53 target genes. *Oncogene* **36**, 3943–3956 (2017). doi: 10.1038/onc.2016.502; pmid: 28288132
44. M. M. Sugrue, D. Y. Shin, S. W. Lee, S. A. Aaronson, Wild-type p53 triggers a rapid senescence program in human tumor cells lacking functional p53. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 9648–9653 (1997). doi: 10.1073/pnas.94.18.9648; pmid: 9275177
45. B. Fito-Lopez, M. Salvadores, M.-M. Alvarez, F. Supek, Prevalence, causes and impact of *TP53*-loss phenocopying events in human tumors. *BMC Biol.* **21**, 92 (2023). doi: 10.1186/s12915-023-01595-1; pmid: 37095494
46. O. Karni-Schmidt, M. Lokshin, C. Prives, The roles of MDM2 and MDMX in cancer. *Annu. Rev. Pathol.* **11**, 617–644 (2016). doi: 10.1146/annurev-pathol-012414-040349; pmid: 27022975
47. M. A. Horlbeck *et al*., Compact and highly active next-generation libraries for CRISPR-mediated gene repression and activation. *eLife* **5**, e19760 (2016). doi: 10.7554/eLife.19760; pmid: 27661255
48. V. Girish, J. M. Sheltzer, A CRISPR competition assay to identify cancer genetic dependencies. *Bio Protoc.* **10**, e3682 (2020). doi: 10.21769/BioProtoc.3682; pmid: 33659353
49. J. I.-J. Leu, P. Dumont, M. Hafey, M. E. Murphy, D. L. George, Mitochondrial p53 activates Bak and causes disruption of a Bak-Mcl1 complex. *Nat. Cell Biol.* **6**, 443–450 (2004). doi: 10.1038/ncb1123; pmid: 15077116
50. M. Mani *et al*., BCL9 promotes tumor progression by conferring enhanced proliferative, metastatic, and angiogenic properties to cancer cells. *Cancer Res.* **69**, 7577–7586 (2009). doi: 10.1158/0008-5472.CAN-09-0773; pmid: 19738061
51. R. Nusse, H. Clevers, Wnt/β-catenin signaling, disease, and emerging therapeutic modalities. *Cell* **169**, 985–999 (2017). doi: 10.1016/j.cell.2017.05.016; pmid: 28575679
52. A. R. Van Rompay, A. Norda, K. Lindén, M. Johansson, A. Karlsson, Phosphorylation of uridine and cytidine nucleoside analogs by two human uridine-cytidine kinases. *Mol. Pharmacol.* **59**, 1181–1186 (2001). doi: 10.1124/mol.59.5.1181; pmid: 11306702
53. D. Sarkisjan *et al*., The cytidine analog fluorocyclopentenylcytosine (RX-3117) is activated by uridine-cytidine kinase 2. *PLOS ONE* **11**, e0162901 (2016). doi: 10.1371/journal.pone.0162901; pmid: 27612203
54. A. Lin *et al*., Off-target toxicity is a common mechanism of action of cancer drugs undergoing clinical trials. *Sci. Transl. Med.* **11**, eaaw8412 (2019). doi: 10.1126/scitranslmed.aaw8412; pmid: 31511426
55. A. Lin, J. M. Sheltzer, Discovering and validating cancer genetic dependencies: Approaches and pitfalls. *Nat. Rev. Genet.* **21**, 671–682 (2020). doi: 10.1038/s41576-020-0247-7; pmid: 32561862
56. A. Luna *et al*., CellMiner Cross-Database (CellMinerCDB) version 1.2: Exploration of patient-derived cancer cell line pharmacogenomics. *Nucleic Acids Res.* **49**, D1083–D1093 (2021). doi: 10.1093/nar/gkaa968; pmid: 33196823
57. S. M. Corsello *et al*., Discovering the anti-cancer potential of non-oncology drugs by systematic viability profiling. *Nat. Cancer* **1**, 235–248 (2020). doi: 10.1038/s43018-019-0018-6; pmid: 32613204
58. E. F. O'Donnell 3rd, H. S. Jang, D. F. Liefwalker, N. I. Kerkvliet, S. K. Kolluri, Discovery and mechanistic characterization of a select modulator of AhR-regulated transcription (SMAhRT) with anti-cancer effects. *Apoptosis* **26**, 307–322 (2021). doi: 10.1007/s10495-021-01666-0; pmid: 33893898
59. Materials and methods are available as supplementary materials.
60. T. Santarius, J. Shipley, D. Brewer, M. R. Stratton, C. S. Cooper, A census of amplified and overexpressed human cancer genes. *Nat. Rev. Cancer* **10**, 59–64 (2010). doi: 10.1038/nrc2771; pmid: 20029424
61. J. Hüllein *et al*., MDM4 is targeted by 1q gain and drives disease in Burkitt lymphoma. *Cancer Res.* **79**, 3125–3138 (2019). doi: 10.1158/0008-5472.CAN-18-3438; pmid: 31000522
62. D. Danovi *et al*., Amplification of *Mdmx* (or *Mdm4*) directly contributes to tumor formation by inhibiting p53 tumor suppressor activity. *Mol. Cell. Biol.* **24**, 5835–5843 (2004). doi: 10.1128/MCB.24.13.5835-5843.2004; pmid: 15199139
63. M. Sebert *et al*., Clonal hematopoiesis driven by chromosome 1q/*MDM4* trisomy defines a canonical route toward leukemia in Fanconi anemia. *Cell Stem Cell* **30**, 153–170.e9 (2023). doi: 10.1016/j.stem.2023.01.006; pmid: 36736290
64. R. J. Quinton *et al*., Whole-genome doubling confers unique genetic vulnerabilities on tumour cells. *Nature* **590**, 492–497 (2021). doi: 10.1038/s41586-020-03133-3; pmid: 33505027
65. Y. Cohen-Sharir *et al*., Aneuploidy renders cancer cells vulnerable to mitotic checkpoint inhibition. *Nature* **590**, 486–491 (2021). doi: 10.1038/s41586-020-03114-6; pmid: 33505028
66. N. Salgia *et al*., Activity of RX-3117, an oral antimetabolite nucleoside, in subjects with advanced urothelial cancer: Preliminary results of a phase IIa study. *J. Clin. Oncol.* **37** (suppl.), 455 (2019). doi: 10.1200/JCO.2019.37.7_suppl.455
67. B. El Hassouni *et al*., Uridine cytidine kinase 2 as a potential biomarker for treatment with RX-3117 in pancreatic cancer. *Anticancer Res.* **39**, 3609–3614 (2019). doi: 10.21873/anticanres.13508; pmid: 31262886
68. C. J. Giuliano, A. Lin, V. Girish, J. M. Sheltzer, Generating single cell-derived knockout clones in mammalian cells with CRISPR/Cas9. *Curr. Protoc. Mol. Biol.* **128**, e100 (2019). doi: 10.1002/cpmb.100; pmid: 31503414

ACKNOWLEDGMENTS

We are grateful to T. Cantz (Hannover Medical School) for providing the TK plasmids used in this work. We thank P. Andrews [Cold Spring Harbor Laboratory (CSHL)] for assistance with SMASH-Seq. We thank Yale Flow Cytometry, especially C. Wang and L. Devine, for their assistance with single-cell sorting. We thank A. Mennone and the Yale Center for Advanced Light Microscopy Facility for assistance with soft agar imaging. We thank the Yale Center for Genome Analysis for performing SMASH-Seq. We thank the Yale Center for Research Computing, specifically R. Bjornson, for guidance and assistance in computations run on the Farnam and Ruddle clusters. We thank the Yale Animal Resources Center Staff for assistance with mouse experiments. Copy number timing analysis conducted in the Sun lab uses the computing resources of the Minnesota Supercomputing Institute. The timing analysis was prepared using limited-access datasets obtained from the Cancer Genome Project from the Wellcome Sanger Institute and does not necessarily reflect the opinions of the provider institution. Part of the BRCA sequencing data was originally generated by research led by M. Kawazu. We also thank the International Cancer Genome Consortium (ICGC) for providing access to the MEL dataset. **Funding:** Research in the Sheltzer lab is supported by NIH grants R01CA237652 and R01CA276666, Department of Defense grant W81XWH-20-1-068, an American Cancer Society Research Scholar Grant, a Breast Cancer Alliance Young Investigator Award, a Damon Runyon-Rachleff Innovation Award, a sponsored research agreement from Ono Pharmaceuticals, and a sponsored research agreement from Meliora Therapeutics. Research in the Liu lab is supported by NIH grant R01GM137031. Yale Flow Cytometry is supported in part by an NCI Cancer Center Support Grant no. NIH P30 CA016359. This work was performed with assistance from the CSHL Flow Cytometry, Microscopy, Animal, and Sequencing Technologies and Analysis Shared Resources, which are supported in part by Cancer Center Support Grant 5P30CA045508. **Author contributions:** Conceptualization: V.G., A.A.L., C.M.S., and J.M.S. Methodology: V.G., A.A.L., C.M.S., S.L.T., A.M.T., Y.L., and J.M.S. Software: R.A.H., J.C.S., and R.S. Formal analysis: R.A.H., J.C.S., and R.S. Investigation: V.G., A.A.L., C.M.S., S.L.T., L.M.B., R.A.H.,

E.L.S., B.E.M., D.A.L., M.L.Y., P.K.K., E.C.S., S.N.L., K.M.S., S.M.A., A.V., C.Z., B.S., and W.L. Resources: A.M.T. Data curation: R.A.H. and J.C.S. Writing – original draft: V.G., A.A.L., and J.M.S. Writing – review & editing: V.G., A.A.L., and J.M.S. Visualization: V.G., A.A.L., S.L.T., L.M.B., R.A.H., R.S., and J.M.S. Supervision: Y.L., R.A.M., R.S., and J.M.S. Funding acquisition: J.M.S. **Competing interests:** J.C.S. is a cofounder of Meliora Therapeutics, a member of the advisory board of Surface Ventures, and an employee of Google, Inc. This work was performed outside of her affiliation with Google and used no proprietary knowledge or materials from Google. J.M.S. has received consulting fees from Merck, Pfizer, Ono Pharmaceuticals, and Highside Capital Management, is a member of the advisory boards of Tyra Biosciences, BioIO, and the Chemical Probes Portal, and is a cofounder of Meliora Therapeutics. All other authors declare that they have no competing interests. **Data and materials availability:** The parental cell lines used in this work are available from various commercial sources (table S4). The plasmids used to induce chromosome loss have been deposited at Addgene (table S5). The code used to perform the TCGA survival analysis is available at https://github.com/joan-smith/comprehensive-tcga-survival (*31*). The code used to perform the mutual exclusivity analysis is available at github.com/sheltzer-lab/aneuploidy-addictions (*27*). The mass spectrometry data have been deposited to the ProteomeXchange Consortium under PRIDE accession no. PXD037956127. RNA-seq data have been deposited in the Gene Expression Omnibus under accession no. GSE222379.

License information: Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adg4521
Materials and Methods
Supplementary Text
Figs. S1 to S21
Tables S1 to S10
References (*69*–*100*)
MDAR Reproducibility Checklist

Submitted 9 January 2023; accepted 27 June 2023
Published online 6 July 2023
10.1126/science.adg4521

RESEARCH ARTICLE SUMMARY

CELL BIOLOGY

The midnolin-proteasome pathway catches proteins for ubiquitination-independent degradation

Xin Gu†, Christopher Nardone†, Nolan Kamitaki, Aoyue Mao, Stephen J. Elledge*, Michael E. Greenberg*

INTRODUCTION: In mammals, the transcriptional response to growth factor, neuronal, and immune stimuli is mediated by a group of genes called immediate-early genes (IEGs), which encode transcription factors of the *Fos*, *EGR*, and *NR4A* families. IEG proteins are activated stereotypically in virtually all mammalian cells but promote the transcription of late-response genes (LRGs) that are cell-type specific and crucial for the appropriate response to the initial stimulus. The physiological importance of IEGs is underscored by the fact that misregulation of their expression can lead to cancer, immune deficiencies, and neurological disorders. The IEG mRNAs accumulate within minutes after the initial stimulus and, once translated, their proteins are rapidly degraded to allow for a transient burst of protein expression. Although the mechanisms that regulate IEG transcription are well characterized, how IEG proteins are swiftly targeted for destruction has remained mysterious for many years.

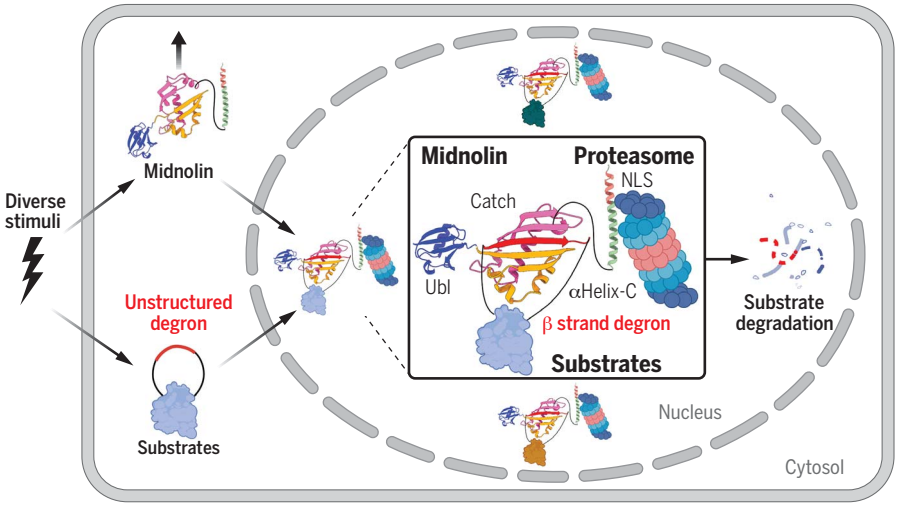
RATIONALE: Eukaryotic cells rely on a macromolecular protease called the proteasome that canonically degrades proteins marked with ubiquitin. It has been suggested that the *Fos* family is targeted to the proteasome by both ubiquitination-dependent and -independent mechanisms, but the molecular events that orchestrate these processes have remained elusive. We hypothesized that there exists a cellular pathway dedicated to the rapid destruction of c-Fos and other IEG proteins. By harnessing the power of forward genetic screens, we sought to identify the machinery that controls the degradation of these proteins.

RESULTS: We performed genome-wide CRISPR-Cas9 screens to search for genes that regulate the stability of IEG proteins. We found that midnolin, a largely uncharacterized protein in mammals, promoted the proteasomal destruction of IEG proteins from structurally distinct families including c-Fos, FosB, EGR1, and NR4A1. These results prompted us to search for additional midnolin targets. We used the global protein stability (GPS) assay with a human open reading frame library (ORFeome) to assess changes in protein stability for ~12,000 human proteins simultaneously. In addition to IEG proteins, midnolin promoted the degradation of IRF4, NeuroD1, PAX8, GATA1, and many other cell-type–specific transcriptional regulators in the nucleus, where midnolin itself resides.

Diverse stimuli that activate IEGs also induced midnolin, and midnolin overexpression was sufficient to cause the destruction of its targets by a mechanism that does not require ubiquitination. Multiple lines of evidence support this ubiquitination-independent mechanism of protein degradation. Midnolin still bound to and promoted the degradation of many targets that had been mutated to lack lysine residues. Moreover, inhibition of the proteasome, but not E1 ubiquitin–activating enzymes, abrogated midnolin function. Additionally, midnolin does not contain RING or HECT domains that are characteristic of E3 ubiquitin ligases or ubiquitin-binding domains found in proteasomal processivity factors such as Rad23.

Instead, midnolin engaged substrates using its "Catch" domain, which was necessary and sufficient to interact with unstructured regions within substrates that have the potential to form a β strand upon binding midnolin. These unstructured regions with the propensity to form a β strand were also necessary and sufficient to bind the Catch domain, thus functioning as a midnolin degron. In addition, midnolin stably associated with the proteasome through a C-terminal α helix and promoted the degradation of Catch-bound targets using its N-terminal ubiquitin-like domain. Thus, midnolin contains three conserved structural domains that function in concert to directly target a large set of nuclear proteins to the proteasome for ubiquitination-independent degradation.

CONCLUSION: Our study suggests that the midnolin-proteasome pathway may represent a general mechanism by which the proteasome bypasses the canonical ubiquitination system to achieve selective degradation of nuclear proteins, many of which are crucial for transcription. Within substrates, midnolin recognizes relatively degenerate amphipathic regions with the potential to form β strands, so the midnolin degron may be a common structural component of numerous proteins. How the midnolin-proteasome pathway is regulated by various cues in diverse cell types to control transcriptional programs will be an important subject of future exploration. ■



The midnolin-proteasome pathway degrades many nuclear proteins independently of ubiquitination. Midnolin is induced by diverse cues, including growth factors and neuronal stimuli. Within the nucleus, midnolin associates with the proteasome through a C-terminal α helix (αHelix-C) and promotes the degradation of bound substrates using an N-terminal ubiquitin-like domain (Ubl). Midnolin achieves selectivity using its Catch domain, which binds an unstructured region within substrates that can form a β strand. Structures shown are AlphaFold models. Figure was created with BioRender.com. NLS, nuclear localization sequence.

The list of author affiliations is available in the full article online.
*Corresponding author. Email: selledge@genetics.med.harvard.edu (S.J.E.); michael_greenberg@hms.harvard.edu (M.E.G.)
†These authors contributed equally to this work.
Cite this article as X. Gu *et al.*, *Science* **381**, eadh5021 (2023). DOI: 10.1126/science.adh5021

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.adh5021

RESEARCH ARTICLE

CELL BIOLOGY

The midnolin-proteasome pathway catches proteins for ubiquitination-independent degradation

Xin Gu[1]†, Christopher Nardone[2,3]†, Nolan Kamitaki[3,4], Aoyue Mao[2,3,5], Stephen J. Elledge[2,3]*, Michael E. Greenberg[1]*

Cells use ubiquitin to mark proteins for proteasomal degradation. Although the proteasome also eliminates proteins that are not ubiquitinated, how this occurs mechanistically is unclear. Here, we found that midnolin promoted the destruction of many nuclear proteins, including transcription factors encoded by the immediate-early genes. Diverse stimuli induced midnolin, and its overexpression was sufficient to cause the degradation of its targets by a mechanism that did not require ubiquitination. Instead, midnolin associated with the proteasome via an α helix, used its Catch domain to bind a region within substrates that can form a β strand, and used a ubiquitin-like domain to promote substrate destruction. Thus, midnolin contains three regions that function in concert to target a large set of nuclear proteins to the proteasome for degradation.

In mammals, extracellular growth factors, cytokines, neurotrophic factors, and neurotransmitters bind their cognate receptors and activate rapid responses by inducing posttranslational modifications of preexisting proteins. More delayed responses are also induced by stimulating gene transcription within the nucleus (*1*, *2*). This transcriptional response occurs in two steps. First, within minutes of the initial stimulus, a set of genes called immediate-early genes (IEGs) is activated. The IEGs encode transcription factors that then trigger a second wave of late-response gene (LRG) expression that mediates the cellular response to the initial stimulus. The IEG response is activated in a stereotypical fashion in virtually every cell type in the body but induces programs of LRG expression that are cell-type specific (*3*). The plethora of cellular responses regulated by IEGs include the cell cycle reentry of quiescent fibroblasts during wound healing; the activation of immune cells in response to cytokines and bacterial and viral infections; and the adaptive responses of neurons to neurotransmitters during learning and memory (*3*).

The IEG family encodes transcription factors such as *Fos*, *Egr*, and *Nr4a* that are rapidly and transiently induced in response to a wide range of extracellular stimuli (*4*). The IEG mRNAs accumulate to a high level, and once these mRNAs are translated, the IEG proteins undergo rapid decay (*5*). Thus, the protein stability of the IEG program is tightly controlled to allow for a relatively brief burst of protein expression that is crucial for appropriate cellular responses to various stimuli. Misregulation of the signaling networks that control IEG expression can lead to cancer, immune deficiencies, and neurological disorders (*3*).

Although the mechanisms that regulate IEG transcription are well characterized, it is unclear how IEG proteins are degraded. In many cases, conjugation of ubiquitin onto substrate proteins occurs as a prelude to their destruction by the proteasome. c-Fos and FosB have been reported to be targeted to the proteasome by both ubiquitination-dependent and -independent mechanisms, but the molecular events that orchestrate these processes are unknown (*6*, *7*).

Genetic screens reveal midnolin as a regulator of IEG protein degradation

To investigate the mechanism of IEG protein degradation, we first used the global protein stability (GPS) reporter system to assay IEG protein stability (*8*). GPS allows for the stable expression of DsRed as an internal control and a green fluorescent protein (GFP)–tagged protein from the same bicistronic mRNA. Therefore, the ratio of GFP/DsRed analyzed by flow cytometry provides a measure of the relative stability of the GFP-fused protein. To identify regulators of IEG protein stability, we generated human embryonic kidney (HEK)–293T cell lines stably expressing the GPS reporter for EGR1 or FosB and performed genome-wide CRISPR-Cas9 screens to search for genes whose disruption stabilized EGR1 or FosB (Fig. 1A). The comparison of EGR1 and FosB allowed us to investigate whether IEGs from different families are degraded by the same or different mechanisms. The top hit from both screens was *MIDN*, a gene that in mammals encodes a largely uncharacterized protein named midnolin (Fig. 1, B and C, and data S1).

To validate our CRISPR-Cas9 screening results, we generated *MIDN* clonal knockout (KO) HEK-293T cells stably NLS expressing the GPS reporter for EGR1 or FosB and assessed the effect of *MIDN* disruption and overexpression on the stability of EGR1 or FosB. Consistent with the screening results, loss of midnolin increased the stability of both EGR1 and FosB (Fig. 1, D and E). We tested additional IEG proteins, and midnolin overexpression was sufficient to decrease the levels of EGR1, FosB, c-Fos, and NR4A1 (Fig. 1, D and E, and fig. S1, A and B) but not several other transcription factors, including ATF2, CREB3, and CREB5 (fig. S1C). These findings raised an interesting possibility that three distinct families of IEG proteins may be targeted for degradation by the same protein: midnolin.

Midnolin is induced and promotes the degradation of several IEG proteins in physiological settings

To investigate the requirement of midnolin for the degradation of IEG proteins in physiologically relevant settings, we mutated *MIDN* using CRISPR-Cas9 to generate a population-level KO of NIH/3T3 fibroblasts, a well-characterized cell line for studying IEG inducibility during cell cycle reentry. Serum deprivation synchronizes NIH/3T3 cells in the G_0 phase of the cell cycle, and IEGs are rapidly and transiently transcribed within minutes after serum addition to these cells (*9*). This transient induction of IEG transcription is followed by an increase in the levels of IEG proteins that then quickly return to their basal levels. In *MIDN* KO cells, IEG protein levels remained high for several hours longer than in control cells, suggesting that the degradation of EGR1, NR4A1, FosB, and c-Fos was attenuated in the absence of midnolin (Fig. 2A). By contrast, stable overexpression of midnolin led to a decrease in the levels of these IEG proteins (Fig. 2B).

In addition to their importance during cell cycle reentry, IEGs mediate adaptive responses in neurons. In response to sensory stimuli, if enough glutamate is released at excitatory synapses in the brain to generate an action potential in the postsynaptic neuron, the depolarization of this neuron results in an influx of calcium that triggers the induction of IEGs (*3*). This membrane depolarization induction of IEGs also occurs in cultured mouse neurons within minutes of exposure to elevated levels of KCl (*10*). To determine whether midnolin regulates IEG protein levels in this paradigm, embryonic mouse cortical neurons were cultured and infected with lentivirus to generate population-level *MIDN* KO or midnolin-overexpressing cells. When exposed to elevated levels of KCl, IEG protein

[1]Department of Neurobiology, Harvard Medical School, Boston, MA 02115, USA. [2]Division of Genetics, Department of Medicine, Howard Hughes Medical Institute, Brigham and Women's Hospital, Boston, MA 02115, USA. [3]Department of Genetics, Harvard Medical School, Boston, MA 02115, USA. [4]Department of Biomedical Informatics, Harvard Medical School, Boston, MA 02115, USA. [5]Department of Molecular and Cellular Biology, Harvard University, Cambridge, MA 02138, USA.
*Corresponding author. Email: selledge@genetics.med.harvard.edu (S.J.E.); michael_greenberg@hms.harvard.edu (M.E.G.)
†These authors contributed equally to this work.

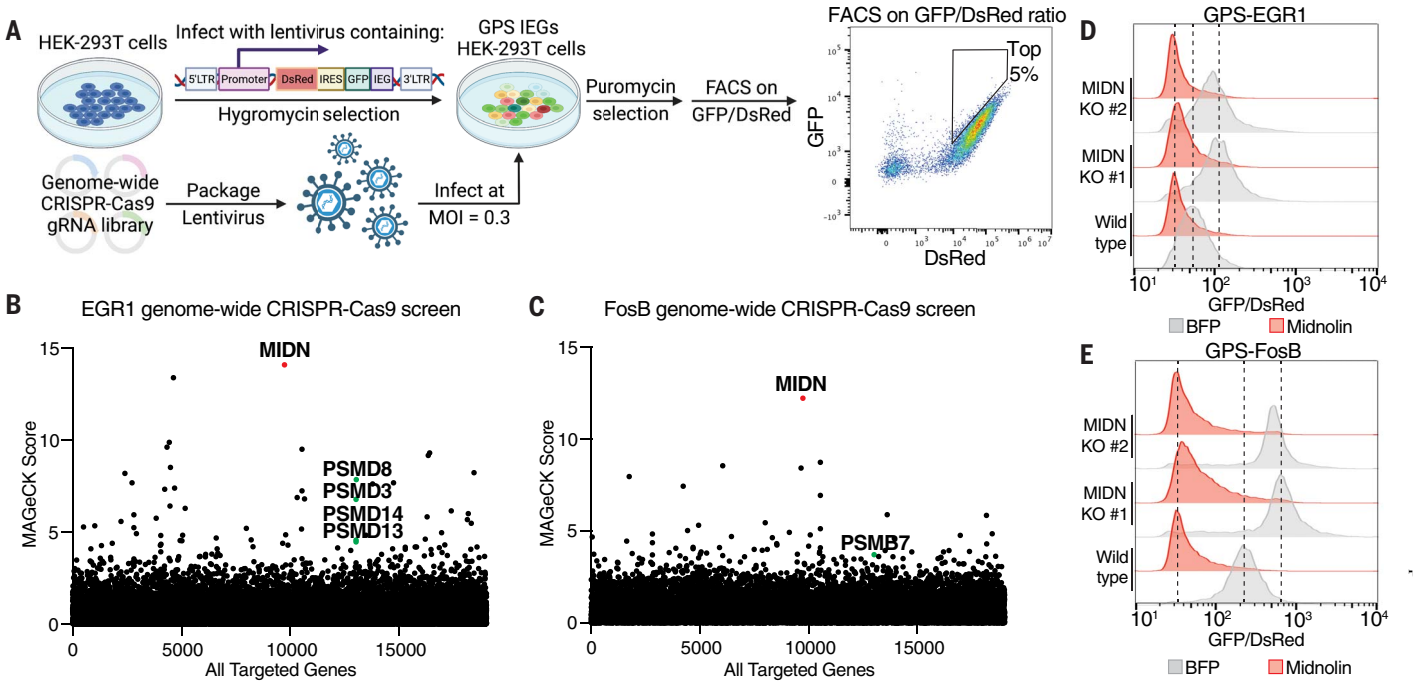


Fig. 1. Genetic screens reveal midnolin as a regulator of IEG protein degradation. (**A**) Schematic showing the FACS-based genome-wide CRISPR-Cas9 screens using the GPS reporter of IEG proteins in HEK-293T cells (figure created with BioRender.com). (**B** and **C**) Results of the genetic screens revealed *MIDN* as the top hit for negatively regulating the stability of both EGR1 and FosB. The proteasomal components showed a weaker enrichment, likely due to their essentiality. The MAGeCK score represents the negative log10 of the "pos|score" value generated from MAGeCK (*62*). (**D** and **E**) Losing midnolin stabilized, whereas overexpressing midnolin destabilized, EGR1 and FosB. GPS EGR1 or FosB reporters were stably expressed in wild-type and two independent *MIDN* KO HEK-293T single cell clones. Vectors expressing BFP control alone (gray) or midnolin and BFP from a EF1α promoter (red) were transiently reintroduced by transfection before analyzing the GFP/DsRed ratio by flow cytometry.

expression was increased in *MIDN* KO neurons but decreased in the midnolin-overexpressing neurons (Fig. 2, C and D). By contrast, IEG mRNA expression was largely unchanged by midnolin overexpression (fig. S1D), which is consistent with midnolin affecting the stability of IEG proteins but not their mRNA transcripts.

KCl treatment of primary cortical neurons induced *MIDN* mRNA levels with kinetics similar to that of IEGs (Fig. 2E). In a previous RNA-sequencing dataset, *MIDN* was found to be induced upon light stimulation of the visual cortex in vivo (fig. S1E) (*11*). Midnolin was also induced upon serum restimulation of NIH/3T3 cells (Fig. 2F) with kinetics similar to that of IEGs (*12*). These stimulus-dependent increases in midnolin expression may be involved in the rapid degradation of IEG proteins. Thus, midnolin is induced by various stimuli and promotes the degradation of IEG proteins in physiologically relevant settings, potentially through a feedback mechanism.

Midnolin can promote the degradation of numerous transcriptional regulators

To determine the extent to which midnolin regulates other cellular proteins beyond IEGs, we performed a screen to identify additional midnolin targets. The screen used a previously described GPS ORFeome library, which contains ~12,000 barcoded human open reading frames (ORFs) tagged with GFP in the GPS reporter system (*13*). The GPS ORFeome library was stably introduced into *MIDN* KO HEK-293T cells, and plasmids expressing either a control blue fluorescent protein (BFP) or midnolin together with BFP were transiently transfected into cells to yield two cell libraries, one lacking midnolin and the other overexpressing midnolin. Because the GPS system overexpresses proteins and endogenous midnolin levels are low, we overproduced midnolin to gain sensitivity in this setting. The cells in each library were then partitioned into six populations on the basis of their GFP/DsRed ratios by fluorescence-activated cell sorting (FACS). The barcodes present in each population were then sequenced to determine the change in their distribution within the cell populations upon midnolin overexpression. If midnolin promoted the destruction of a given barcoded GFP-fusion protein, then the distribution of the barcode would shift to a cell population with a lower GFP/DsRed ratio in the midnolin-overexpressing library (Fig. 3A). This screen yielded our previously characterized targets of midnolin, including FosB and c-Fos, along with CBX4 (Pc in flies), which was previously shown to be regulated by the *Drosophila melanogaster* ortholog of midnolin, Stuxnet (*14*) (Fig. 3B). Midnolin overexpression also led to a robust reduction in the GFP/DsRed ratios of many proteins, consistent with a decrease in protein stability (data S2). Most of the proteins regulated by midnolin overexpression were nuclear proteins that regulate transcription (Fig. 3C), such as the lineage-specific transcription factors IRF4, NeuroD1, PAX8, and GATA1.

To validate the findings from the GPS ORFeome screen, midnolin was overexpressed in *MIDN* KO HEK-293T cells stably expressing individual GPS reporters for the identified proteins. Midnolin overexpression was sufficient to reduce the GFP/DsRed ratios of these proteins, consistent with a decrease in their stability (Fig. 3D and fig. S2A). The endogenous levels of various proteins identified in the screen were also substantially reduced in HEK-293T cells that expressed a doxycycline-induced midnolin (Fig. 3E). Among the lineage-specific transcription factors that are not expressed in HEK-293T cells, interferon regulatory factor 4 (IRF4) is essential for the function and homeostasis of mature B and T cell lymphocytes (*15*, *16*). To test the ability of midnolin to promote the destruction of endogenous IRF4, we generated Ramos B cell lines with population-level KO or overexpression of midnolin. In these cell lines, the steady-state levels of IRF4 were markedly increased in the *MIDN* KO cells and decreased in the midnolin-overexpressed cells (Fig. 3F). Thus,

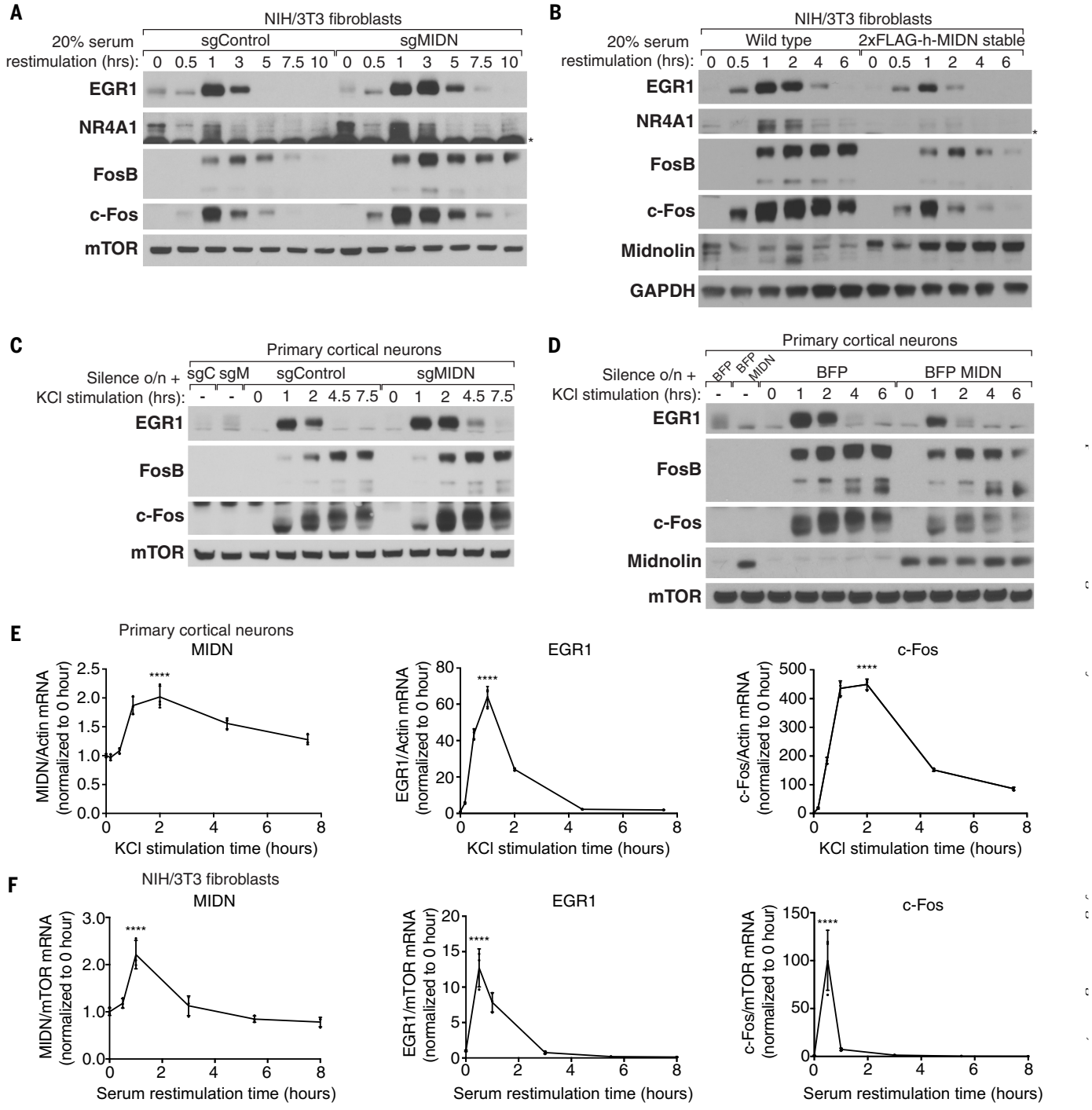


Fig. 2. Midnolin is induced and promotes the degradation of several IEG proteins in physiological settings. (**A**) Loss of midnolin increased the expression of IEG proteins in NIH/3T3 cells. Immunoblotting was performed from NIH/3T3 cells stably expressing Cas9 and control or *MIDN* targeting sgRNAs. This population-level mutagenesis of *MIDN* may show lower penetrance relative to an isogenic mutant because the KO efficiency is dependent on the efficacy of the sgRNA. The cells were starved of serum overnight before serum restimulation for the indicated time points. Asterisks mark nonspecific cross-reactive proteins. (**B**) Overexpressing midnolin decreased the expression of IEG proteins in NIH/3T3 cells. Same assay as in (A) but in NIH/3T3 cells stably overexpressing an N-terminally 2xFLAG tagged human midnolin using a CMV promoter. (**C**) Loss of midnolin increased the expression of IEG proteins in primary cortical neurons. Neurons were isolated from E16.5 mouse brains and cultured in a dish. On day 3 after isolation, the neurons were infected with lentivirus encoding Cas9 with control or *MIDN* targeting sgRNAs. Immunoblotting was performed on day 11 after dissection from neurons that were silenced overnight with TTX (a sodium channel blocker) and D-AP5 (an NMDA receptor antagonist) and stimulated with KCl for the indicated time points to induce depolarization. (**D**) Overexpressing midnolin decreased the expression of IEG proteins in primary cortical neurons. Similar assay as in (C) but using lentivirus to overexpress a BFP control or human midnolin coexpressing BFP using an EF1α promoter. (**E** and **F**) qPCR analysis for mRNA levels of the indicated genes from primary mouse cortical neurons that were KCl stimulated (E) or from NIH/3T3 cells that were serum restimulated (F) for the indicated time points. Error bars indicate SD from three biological replicates. Data were analyzed using an ordinary one-way ANOVA followed by Tukey's multiple-comparisons test; ****$P < 0.0001$.

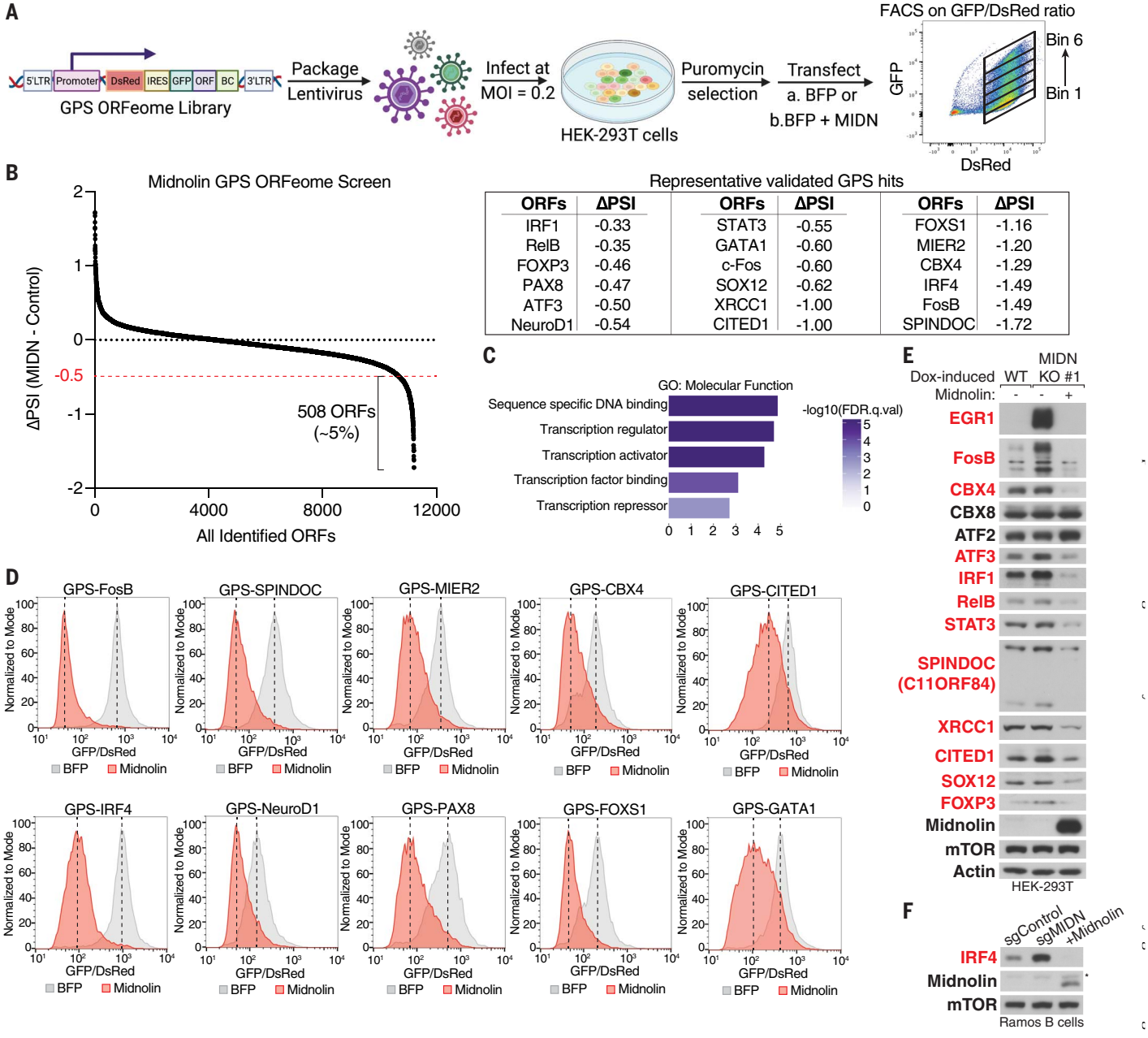


Fig. 3. Midnolin can promote the degradation of numerous transcriptional regulators. (**A**) Schematic showing the midnolin GPS ORFeome screen. The GPS ORFeome library (~12,000 barcoded human ORFs tagged to GFP) was introduced into *MIDN* KO HEK-293T, and the library-expressing cells were transfected with BFP control or midnolin coexpressing BFP before FACS sorting the library into populations based on the GFP/DsRed ratio (figure created with BioRender.com). (**B**) Analysis of the GPS ORFeome screen showing the change in protein stability (ΔPSI) between midnolin and BFP, which was calculated based on the change in read distribution of the barcoded ORFs. Approximately 5% of the library showed significant destabilization with ΔPSI values <−0.5. Several validated hits from the screen are shown in the boxed table. (**C**) GSEA based on the GPS ORFeome screen for molecular function. (**D**) Validation of screen hits indicates their potent regulation by midnolin. GPS reporters for the indicated genes were stably expressed in *MIDN* KO HEK-293T cells, and a control BFP or midnolin coexpressing BFP was transiently transfected before analyzing the GFP/DsRed ratio by flow cytometry. (**E**) Endogenous proteins of numerous screen hits are regulated by midnolin. Immunoblotting was performed from wild-type, *MIDN* KO, and *MIDN* KO HEK-293T cells in which midnolin expression was stably induced with doxycycline (100 ng/ml) for 2 days using a TRE promoter. Shown are putative midnolin targets (red) based on the GPS ORFeome screen and negative controls (black). (**F**) Validation of midnolin-mediated degradation of endogenous IRF4 in Ramos B cells. Immunoblotting was performed from Ramos B cells expressing Cas9, control or *MIDN* targeting sgRNAs, or stably overexpressing midnolin using an EF1α promoter.

through gain-of-function screening, we uncovered many potential targets of midnolin that have important tissue or cell type–specific functions in regulating gene expression.

Midnolin associates both with its substrates and the 26*S* proteasome

To begin to determine how midnolin promotes the degradation of a large diverse set of proteins, we generated a HEK-293T cell line in which endogenous midnolin was tagged at its N terminus with 3xHA to facilitate immunoprecipitation. We found by mass spectrometry

that endogenous midnolin coimmunoprecipitated essentially all proteasomal subunits of the 19*S* regulatory particle and the 20*S* proteolytic core particle (Fig. 4A and data S3). These proteasomal subunits were the most abundant proteins detected by mass spectrometry. Also, there were no proteins aside from the proteasomal proteins that were both coimmunoprecipitated by midnolin and scored in our genome-wide CRISPR-Cas9 screens (data S3 and Fig. 1, A to C). These results suggested that midnolin likely interacts directly with the proteasome. In addition, by mining the BioPlex protein-protein interaction dataset (*17*), we found that several IEG proteins, including EGR1, FosB, and NR4A1, as well as a proteasomal component, PSMD2, coimmunoprecipitated endogenous midnolin in HEK-293T cells (fig. S3A).

To confirm the mass spectrometry findings, we treated HEK-293T cells expressing 3xHA-tagged midnolin with MG132 to prevent substrate degradation or with phorbol 12-myristate 13-acetate (PMA), a protein kinase C agonist, to induce the transcription of IEGs (*18*). Endogenous midnolin interacted with c-Fos, FosB, EGR1, and NR4A1, as well as the proteasome, as indicated by PSMD2 and PSMA2, components of the 19*S* and 20*S* proteasome, respectively (Fig. 4B). Like serum-stimulated fibroblasts and KCl-treated neurons, exposure of HEK-293T cells to PMA led to an increase in the level of the midnolin protein (Fig. 4B). Thus, midnolin interacts with the proteasome to promote the degradation of midnolin-bound proteins.

Midnolin promotes ubiquitination-independent degradation of bound substrates

We next investigated whether midnolin targets its substrates for destruction by a ubiquitination- and proteasome-dependent mechanism. Individual GPS reporters of midnolin substrates were stably expressed in *MIDN* KO HEK-293T cells, which were transfected with BFP control or midnolin coexpressing BFP. These reporter cells were then treated for 6 hours with the proteasome inhibitor MG132 or TAK-243, a potent inhibitor of the E1 ubiquitin–activating enzymes UBA6 and UAE that inhibits protein ubiquitination globally (*19*). Although the proteasome inhibitor MG132 strongly reduced the midnolin-mediated degradation of these midnolin substrates, the ubiquitin E1 inhibitor TAK-243 did not disrupt midnolin function (fig. S3B). By contrast, both MG132 and TAK-243 promoted stabilization of c-Myc, which does not appear to be a midnolin substrate (fig. S3B). Furthermore, when we treated HEK-293T cells expressing 3xHA-tagged midnolin with MG132 or TAK-243 for 6 hours, we found that the proteasome inhibitor MG132 led to a potent increase in the expression of the midnolin protein, but the E1 inhibitor TAK-243 caused a slight decrease in the level of midnolin. Conversely, the level of c-Myc and p27, two transcriptional regulators that are known to be targeted for proteasomal degradation in a ubiquitination-dependent manner, increased upon exposure to TAK-243 or MG132 (Fig. 4C) (*20*–*22*). Thus, midnolin does not appear to require ubiquitination for its own turnover by the proteasome.

To further test the requirement of ubiquitination for midnolin-mediated degradation, potential ubiquitination sites were mutated in several midnolin substrates. Because canonical ubiquitination occurs on lysine residues, all lysine residues within these substrates were mutagenized to arginine to block lysine-dependent ubiquitination (*23*–*25*). Wild-type and K-to-R mutant substrates, such as EGR1, FosB, c-Fos, NR4A1, NeuroD1, and IRF4, interacted with endogenous midnolin to a similar extent (Fig. 4D). When stably expressed in *MIDN* KO HEK-293T cells, both wild-type and K-to-R mutant substrates were efficiently degraded upon doxycycline-induced midnolin overexpression (Fig. 4E), indicating that lysine residues were not required for midnolin-dependent substrate destruction. Thus, it seems that midnolin directly associates with the proteasome and promotes the degradation of many transcriptional regulators without requiring their ubiquitination.

Midnolin contains three domains that function in concert to promote proteasomal degradation of bound substrates

To gain insight into how midnolin interacts with the proteasome and its numerous substrates, we used AlphaFold to obtain a predicted structure of midnolin, which revealed three confidently predicted and highly conserved regions with defined structure (Fig. 5A and fig. S4, A and B) (*26*). Midnolin does not contain structural elements that are characteristic of RING- or HECT-type ubiquitin ligases (*27*) and does not contain a ubiquitin-binding domain characteristic of proteasomal processivity factors such as Rad23 or ubiquilin (*28*). Instead, midnolin contains a ubiquitin-like domain (Ubl) toward its N terminus. Additionally, midnolin contains two discontinuous regions, each composed of two predicted antiparallel β strands and two or three α helices, that appear to fold together to form a domain with internal symmetry. For reasons discussed below, we named this region of midnolin the "Catch" domain. Finally, midnolin contains a long α helix toward its C terminus, called αHelix-C, which includes a predicted nuclear localization sequence (NLS). Indeed, endogenous midnolin was largely located within the nucleus, and deletion of the predicted NLS, but not other regions of midnolin, resulted in its localization to the cytoplasm (fig. S5, A and B).

To determine whether these three regions are important for midnolin function, we transiently expressed wild-type and mutant versions of midnolin in *MIDN* KO HEK-293T cells stably expressing the GPS IRF4 or FosB reporters (Fig. 5, B and C). Although wild-type midnolin potently promoted IRF4 and FosB destruction, point mutations of the Ubl or deletions of the Ubl, Catch, αHelix-C, or NLS domains abrogated midnolin function (Fig. 5C and fig. S6A).

We next performed coimmunoprecipitation experiments to identify midnolin domains that are required for its interaction with substrates and/or the proteasome. Point mutations or deletion of the Ubl domain did not disrupt the stable association of midnolin with EGR1 or the proteasome (Fig. 5D). However, mutagenesis of the Ubl domain potently increased midnolin levels, and these Ubl mutants were only marginally sensitive to MG132 (Fig. 5D). This suggests that the ubiquitination-independent degradation of midnolin that we observed previously (Fig. 4C) is dependent on its Ubl domain. By contrast, deletion of the NLS or entire C-terminal α helix revealed that this domain is necessary for midnolin to interact stably with the proteasome but not its substrates (Fig. 5D). This association with the proteasome was not affected by inhibition of the proteasome or E1 ubiquitin–activating enzymes (fig. S6B). When fused to maltose-binding protein (MBP), the midnolin helix conferred the ability to interact with the proteasome (Fig. 5E). Thus, the C-terminal midnolin helix is both necessary and sufficient to bind the proteasome and midnolin and, unlike the processivity factors Rad23 and ubiquilin, engages the proteasome stably independently of the Ubl domain (Fig. 5D) (*29*, *30*).

Deletion of the regions that fold together to form the Catch domain (the N-terminal Catch1 and C-terminal Catch2 subdomains) abolished the interaction of midnolin with its substrates without affecting its ability to bind the proteasome (Fig. 5D). In growth-arrested NIH/3T3 cells in which the level of endogenous midnolin substrates was induced upon serum restimulation, the interaction of midnolin with its substrates also required the Catch domain (fig. S6C). To determine whether the Catch domain is both necessary and sufficient to engage midnolin substrates, we immunoprecipitated transfected wild-type, Catch domain–deleted midnolin or the Catch domain alone in HEK-293T cells and assessed the interaction with various substrates and proteasomal components. Deletion of the Catch domain abolished the interaction of midnolin with its substrates while retaining proteasome binding, and the Catch domain alone was sufficient to bind midnolin substrates (Fig. 5F). Catch1 and Catch2 are separated by a long, 111–amino acid unstructured sequence that is poorly conserved. We shortened this Catch1 and Catch2 linker to 10 amino acids (ΔLoop1), 16 amino acids (ΔLoop2), or 28 amino acids (ΔLoop3). These midnolin variants still interacted with substrates and promoted their degradation, indicating that the long stretch connecting Catch1 and Catch2 is largely dispensable (Fig. 5F and fig. S7, A and B). However, the length between Catch1 and Catch2 could not be too short,

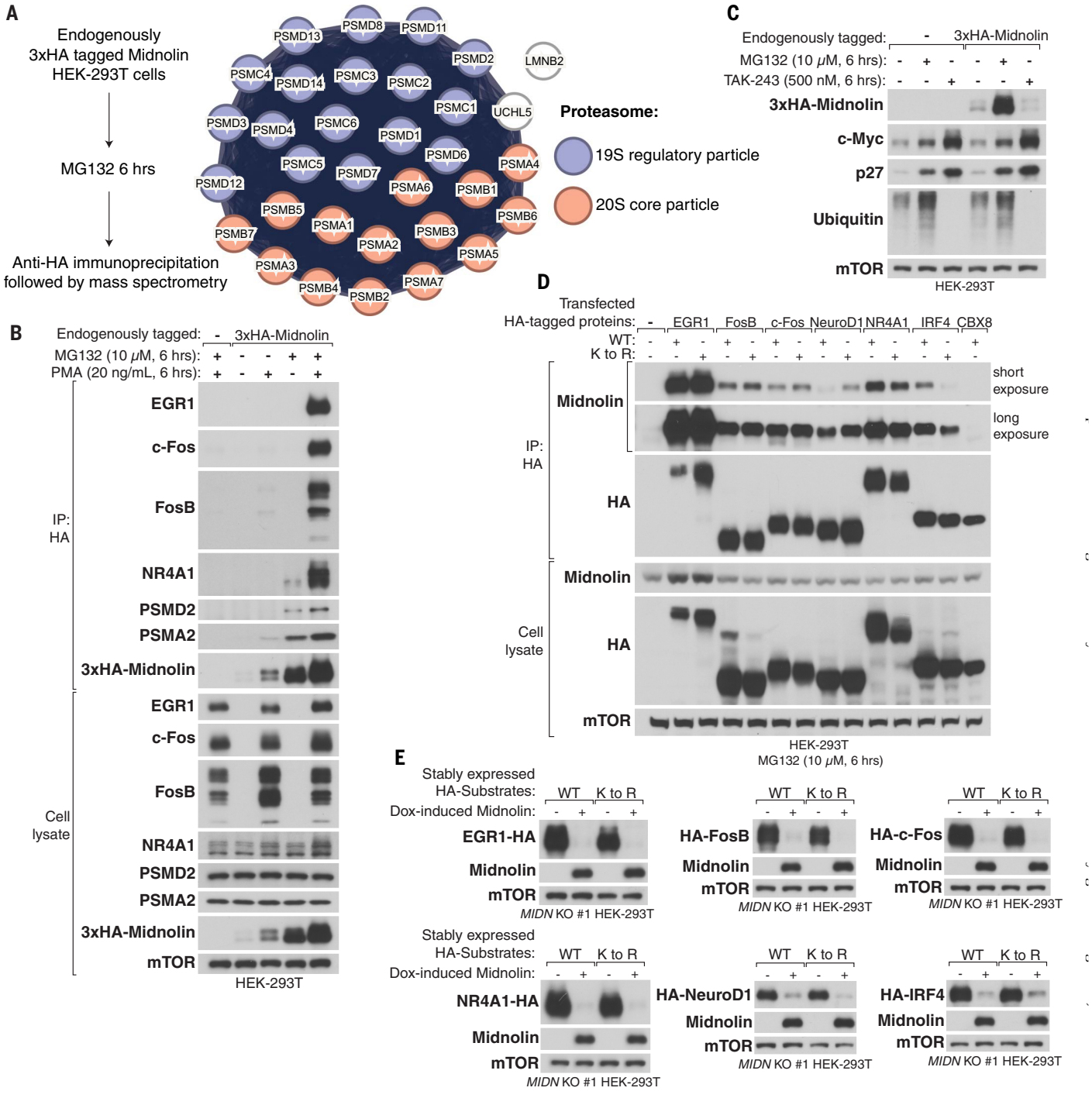


Fig. 4. Midnolin associates with the proteasome to promote ubiquitination-independent degradation of bound substrates. (**A**) A 3xHA-tag was introduced at the N terminus of the endogenous midnolin locus in HEK-293T cells using CRISPR-Cas9–initiated recombination. Cells were treated with MG132 for 6 hours before immunoprecipitation of 3xHA-midnolin followed by mass spectrometry. The results revealed a large enrichment of the 26*S* proteasome (data S3), and shown is a STRING analysis of the top coimmunoprecipitated proteins identified from the mass spectrometry. (**B**) Midnolin coimmunoprecipitates the proteasome and IEG proteins endogenously. Immunoblotting was performed from anti-HA immunoprecipitants of endogenous 3xHA-midnolin from the knock-in HEK-293T cells treated with the indicated drugs for 6 hours. PMA was used to induce the transcription of IEGs. (**C**) Endogenous midnolin protein levels are strongly increased by proteasomal inhibition but not by ubiquitin E1 inhibition. Immunoblotting was performed from wild-type and 3xHA-midnolin knock-in HEK-293T cells treated with 10 μM MG132 or 500 nM TAK-243 for 6 hours. (**D**) Lysine-dependent ubiquitination on substrates is not necessary for midnolin interaction. Immunoblotting was performed from anti-HA immunoprecipitants of HEK-293T cells that were transfected with the indicated constructs, either wild-type or all lysine residues mutated to arginine residues (K-to-R mutants). Cells were treated with 10 μM MG132 for 6 hours. CBX8 serves as a negative control as it is not targeted by midnolin. (**E**) Midnolin does not require lysine residues on substrates to promote degradation. Wild-type and K-to-R mutant substrates were stably introduced into *MIDN* KO HEK-293T cells using a cytomegalovirus (CMV) promoter. Then, midnolin expression was induced using doxycycline (100 ng/ml) for 2 days using a TRE promoter before lysis and immunoblotting.

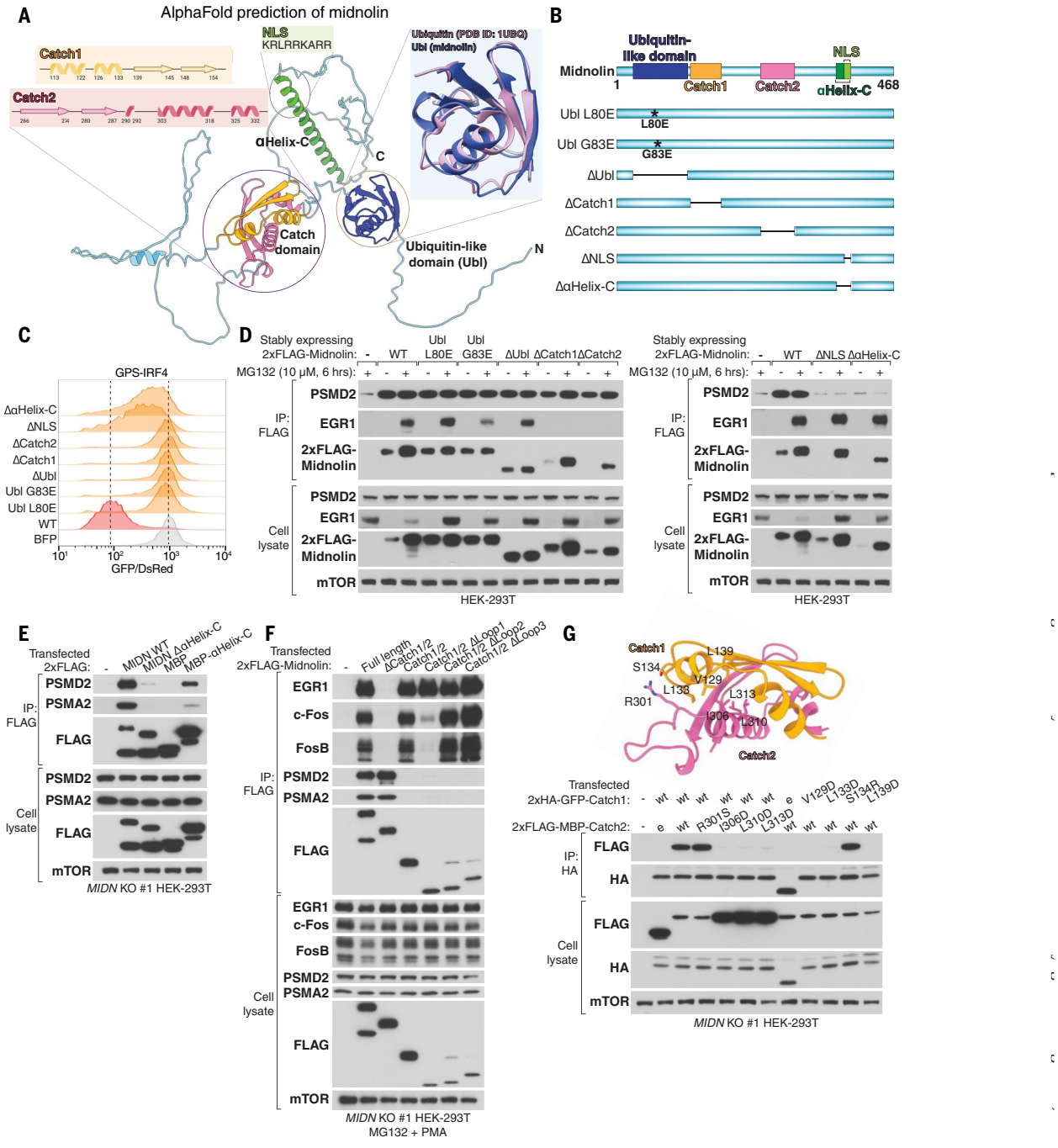


Fig. 5. Midnolin contains three regions that function in concert to promote proteasomal degradation of bound substrates. (**A**) Midnolin structure prediction by AlphaFold (Q504T8-F1) reveals three regions with defined structure (*26*). (**B**) Schematic representation of mutations (*72*) or truncations introduced into the midnolin cDNA. See the materials and methods for the truncation boundaries and regions used for sufficiency experiments. (**C**) Regions with defined structure are necessary for a functional midnolin. The GPS IRF4 reporter was stably expressed in *MIDN* KO HEK-293T cells, and a control BFP or wild-type and mutant versions of midnolin coexpressing BFP were transiently transfected before analyzing the GFP/DsRed ratio by flow cytometry. (**D**) The midnolin Catch domain is necessary for binding substrates, and the C-terminal α helix is necessary for proteasomal association. Immunoblotting was performed from anti-FLAG immunoprecipitants of HEK-293T cells stably expressing 2xFLAG-tagged midnolin using a CMV promoter. Cells were treated with 10 μM MG132 for 6 hours. (**E**) The midnolin αHelix-C is sufficient to interact with the proteasome. Immunoblotting was performed from anti-FLAG immunoprecipitants of *MIDN* KO HEK-293T cells transfected with the indicated 2xFLAG-tagged proteins. (**F**) The midnolin Catch domain is sufficient to bind substrates. Immunoblotting was performed from anti-FLAG immunoprecipitants of *MIDN* KO HEK-293T cells transfected with the indicated 2xFLAG-tagged proteins. The 111–amino acid sequence between Catch1 and Catch2 was shortened to 10 amino acids (ΔLoop1), 16 amino acids (ΔLoop2), or 28 amino acids (ΔLoop3). Cells were treated with 10 μM MG132 and 20 ng/ml of PMA for 6 hours. (**G**) The Catch1 and Catch2 regions of midnolin interact when expressed as independent proteins. Immunoblotting was performed from anti-HA immunoprecipitants of *MIDN* KO HEK-293T cells cotransfected with 2xHA-GFP-Catch1 and 2xFLAG-MBP-Catch2 constructs, where "e" signifies empty 2xHA-GFP or 2xFLAG-MBP.

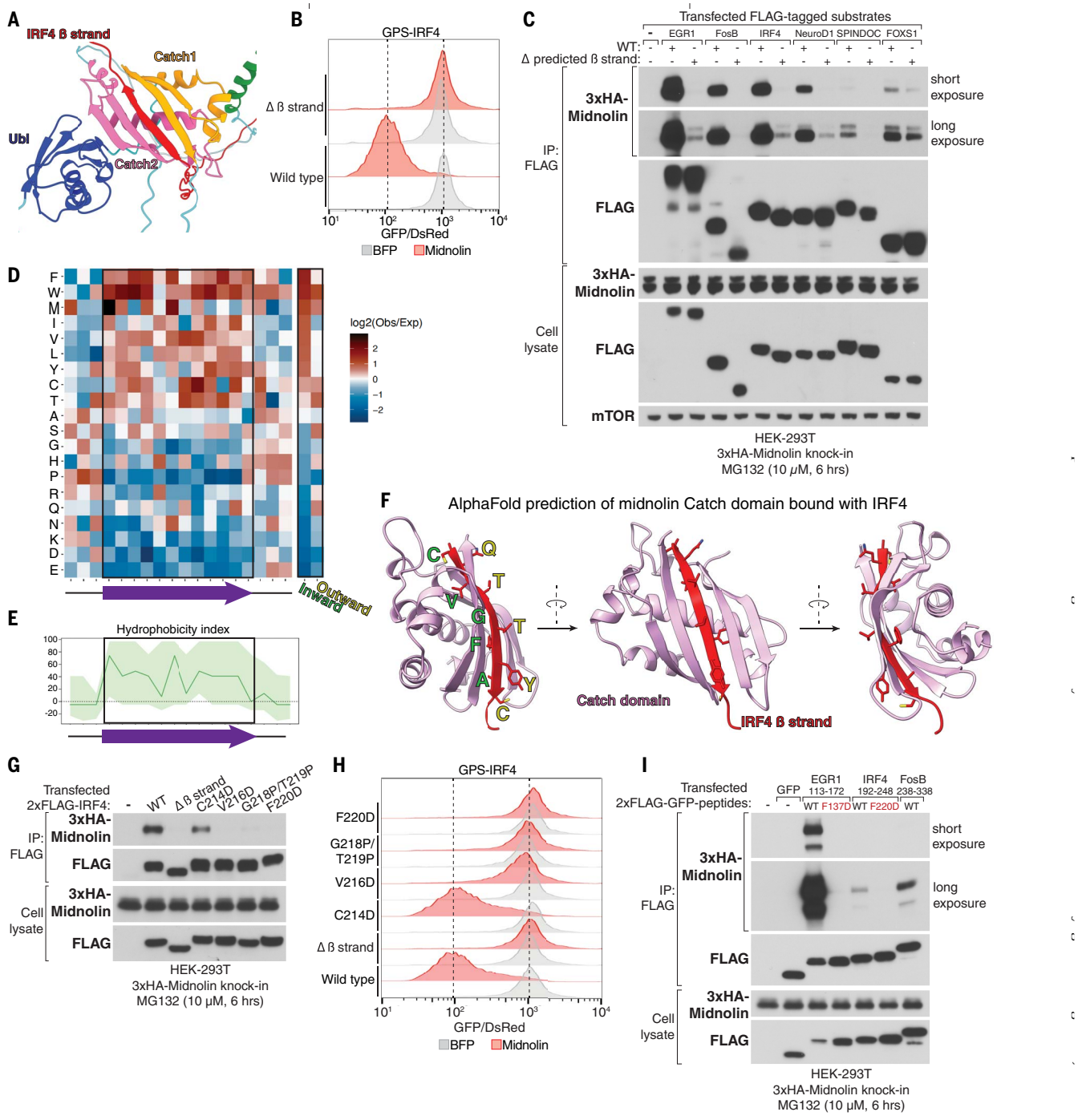


Fig. 6. Midnolin catches regions within its substrates that constitute a β-strand degron. (**A**) AlphaFold structure prediction of midnolin bound to its substrate IRF4 reveals an adopted β-strand capture model. (**B**) Midnolin requires the predicted β strand within IRF4 to promote degradation. The GPS IRF4 reporters were stably expressed in *MIDN* KO HEK-293T cells, and a control BFP or midnolin coexpressing BFP was transiently transfected before analyzing the GFP/DsRed ratio by flow cytometry. (**C**) Predicted β strands are necessary for interaction with midnolin. Immunoblotting was performed from anti-FLAG immunoprecipitants of 3xHA-midnolin knock-in HEK-293Ts transfected with 2xFLAG-tagged substrates. For FosB, the comparison is between the full-length protein and ΔFosB. Cells were treated with 10 μM MG132 for 6 hours. See the materials and methods for the truncation boundaries. (**D**) Amino acid frequency of midnolin substrate β strands predicted by AlphaFold reveals a strong preference for hydrophobic residues. "Inward" is defined by the residues buried within the Catch domain, and "Outward" is defined by the solvent-exposed residues. (**E**) The hydrophobicity of residues within the β strand was determined by a mean hydrophobicity index at pH 7 (*70*, *71*) of residues immediately before, within, or immediately after the β strand. (**F**) AlphaFold structure prediction of the midnolin Catch domain bound to IRF4. (**G**) Hydrophobic β-strand residues buried within the Catch domain are required for midnolin interaction [similar assay as in (C) but from cells transfected with the 2xFLAG-tagged IRF4 constructs]. (**H**) Midnolin requires the hydrophobic β-strand residues buried within the Catch domain to promote degradation [similar assay as in (B)]. (**I**) Regions encompassing predicted β strand(s) are sufficient for conferring an interaction with midnolin [similar assay as in (C) but from cells transfected with the indicated 2xFLAG-GFP-peptide fusions].

because the shortest linker (ΔLoop1, 10 amino acids) did attenuate the interaction of midnolin with certain substrates (Fig. 5F and fig. S7A). Thus, the linker between Catch1 and Catch2 may still be important for proper flexibility or folding of the Catch domain.

Given the importance of the Catch domain for binding substrates, we further validated the AlphaFold prediction that Catch1 and Catch2 fold together. This folding may be strongly driven by hydrophobicity, because the core of the Catch domain is composed exclusively of highly conserved hydrophobic amino acids (fig. S7C). To validate that Catch1 and Catch2 fold back to bind each other in a process driven by hydrophobicity, we expressed each separately by tagging Catch1 to 2xHA-GFP and Catch2 to 2xFLAG-MBP. We also generated a series of mutants for both Catch1 and Catch2 in which some hydrophobic residues predicted by AlphaFold to drive the interaction were changed to aspartic acid. As a control, we mutagenized a solvent-exposed residue on Catch1 (S134) and Catch2 (R301) because these were not predicted to contribute to the binding between Catch1 and Catch2. These epitope-tagged Catch1 and Catch2 constructs were expressed in *MIDN* KO HEK-293T, and 2xHA-GFP-Catch1 was immunoprecipitated. Indeed, Catch1 and Catch2 bound each other when expressed as independent proteins, and mutagenesis of the hydrophobic residues to aspartic acid abolished the interaction (Fig. 5G). This contrasts with mutagenesis of the solvent-exposed residues, which did not alter the Catch1-Catch2 interaction as anticipated. We also introduced these same mutations into full-length midnolin, and only mutagenesis of the hydrophobic residues abolished the ability of midnolin to bind with and promote the degradation of its substrates (fig. S7, B and D). The hydrophobic core was important for the function of the Catch domain, because mutagenesis of additional hydrophobic residues to aspartic acid attenuated the ability of midnolin to promote the degradation of its substrates, including IRF4 and FosB (fig. S7E).

Therefore, midnolin is primarily localized within the nucleus, associates with the proteasome using its long C-terminal α helix, binds substrates using its hydrophobic Catch domain, and contains a Ubl domain that is necessary to promote substrate degradation. Overall, these three regions of midnolin function in concert to promote ubiquitination-independent proteasomal degradation of bound substrates.

Midnolin "catches" regions within its substrates that constitute a β-strand degron

We wondered how midnolin achieves substrate selection through its Catch domain given its ability to promote the destruction of many diverse proteins. Canonically, E3 ubiquitin ligases bind short linear motifs within substrates, called degrons, in which the amino acid side chains play crucial roles in determining substrate specificity (*31*–*34*). To gain insights into how midnolin achieves substrate selection, we used AlphaFold to predict the structure of the midnolin-substrate bound complex (*26*). The AlphaFold predictions for midnolin bound to IRF4 (Fig. 6A) revealed that a predicted unstructured region of native IRF4 (fig. S8A) formed a β strand upon binding to the midnolin Catch domain (fig. S8B), thereby completing a five-stranded antiparallel β-sheet tertiary structure. Consistent with this AlphaFold-predicted interaction, a small deletion within IRF4 that encompassed the predicted β strand abolished the ability of midnolin to promote IRF4 degradation (Fig. 6B).

To determine the generality of this β-strand capture mechanism, we performed the AlphaFold structure predictions for the 508 most destabilized proteins uncovered in the GPS ORFeome screen. The Protein Data Bank (PDB) files for these 508 predictions can be accessed on Dryad using the DOI 10.5061/dryad.m905qfv6g. Approximately 40% (205/508) of the proteins have predictions that are consistent with them being captured by midnolin (data S4). We systematically compared the AlphaFold-predicted structure of midnolin substrates in the native and midnolin-bound states. The regions predicted to be captured by midnolin are generally more unstructured than the rest of the same protein in the native state but have the potential to form a β strand upon binding the Catch domain (fig. S8C). We validated these findings by introducing into several representative midnolin substrates small deletions that encompassed the predicted β strand and tested the ability of midnolin to interact with and promote the degradation of these mutants. In each case, deletion of the predicted β strand abrogated the ability of midnolin to both interact with substrates (Fig. 6C) and promote their degradation (fig. S9, A to D).

The AlphaFold predictions yielded several different modes of midnolin-FosB interaction, raising the possibility that the N and C terminus of FosB have the potential to form multiple different β strands that could be captured by midnolin. We generated various truncated forms of FosB to determine which regions were necessary for midnolin to promote degradation and found that the C-terminal 101 amino acids were required (fig. S10A). Truncation of the last 101 amino acids of FosB results in a naturally existing splice isoform called ΔFosB, which was shown to be significantly more stable than other Fos family members, although the molecular basis for this increased stability was unknown (*35*). Previous studies have shown that chronic and repeated exposure to drugs of abuse such as cocaine leads to week-long accumulation of ΔFosB protein expression in the nucleus accumbens, a brain region crucial for addiction (*35*, *36*). Indeed, overexpression of ΔFosB in neurons of the nucleus accumbens sensitizes animals to the effects of cocaine and may thus contribute to addiction (*37*). In contrast to full-length FosB, midnolin weakly interacted with and less effectively promoted the degradation of ΔFosB (Fig. 6C and fig. S10A). Our findings therefore provide a mechanistic explanation for the increased stability of ΔFosB. Nevertheless, ΔFosB is eventually degraded, and this may require the N terminus of FosB, which is also predicted by AlphaFold to form a β strand that can be captured by midnolin. Deletion of the predicted N-terminal β strand region within ΔFosB largely abolished its ability to be targeted for decay by midnolin (fig. S10B). Thus, a protein can contain more than one region that can be captured by midnolin.

Having identified sequences that mediate the midnolin-substrate interaction, we investigated whether specific amino acids were enriched or depleted over others within the predicted β strands captured by the midnolin Catch domain. Within these β strands, we observed a marked depletion of charged amino acids, including aspartic acid and glutamic acid, as well as proline (Fig. 6D), which is known to disrupt β strands (*38*). Instead, there was a strong enrichment for hydrophobic amino acids within the midnolin-induced β strands, which overall were significantly more hydrophobic than the average of all other regions within the same protein (Fig. 6E and fig. S10C). Once bound by midnolin, these hydrophobic β strand residues were predicted to be buried within the core of the Catch domain (inward), whereas charged amino acids tended to point outward and were solvent exposed (Fig. 6, D and F, and fig. S10D). The enrichment for hydrophobicity within the β strands is consistent with the fact that the core of the Catch domain is also highly hydrophobic and required for catching hydrophobic regions within substrates. To validate these predictions, two residues (G218 and T219) at the center of the predicted IRF4 β strand were mutagenized to proline to potentially disrupt β-strand formation, or the hydrophobic residues buried within the interior of the Catch domain (V216, and F220) were mutagenized to aspartic acid to disrupt potential hydrophobic interactions. Consistent with the AlphaFold predictions, both the introduction of prolines and mutagenesis of the hydrophobic residues to aspartic acid within the β strand abolished the ability of midnolin to interact with (Fig. 6G) and promote the degradation of IRF4 (Fig. 6H). Similar results were obtained after mutagenesis of the corresponding residues within the EGR1-adopted β strand (fig. S10, E and F).

We next tested whether midnolin could interact with and promote the degradation of GFP fused to a short sequence containing the β-strand degron(s) within substrates. Midnolin interacted with (Fig. 6I) and promoted the degradation of GFP fused to short stretches within EGR1, IRF4,

and the C-terminal tail of FosB that was dependent on the hydrophobicity of the predicted β strand (fig. S10, G and H). Thus, the β strands predicted by AlphaFold are necessary and sufficient to interact with midnolin for proteasomal degradation.

We conclude that midnolin achieves substrate selection through a general mechanism: Captured regions within substrates have the ability to form a β strand that is biochemically compatible with the hydrophobic core of the Catch domain.

Discussion

In this study, we identified a protein called midnolin that targets stimulus-induced transcription factors such as c-Fos, FosB, EGR1, NR4A1, and IRF4 to the proteasome for degradation. Disrupting midnolin function in various cell types increases the peak abundance of these transcription factors and prolongs their expression. It is not uncommon for a protein to be targeted for proteasomal degradation by multiple mechanisms, and these stimulus-responsive transcription factors may not be targeted for degradation solely by midnolin. Rather, it is possible that these proteins are also targeted by a ubiquitination-dependent pathway that functions in parallel with midnolin and could be the primary mechanism of degradation in some contexts. Indeed, it has been reported that Fos family members are targeted by both ubiquitination-dependent and -independent mechanisms (*7*). We found that midnolin is induced by diverse stimuli, and its induction may act as a post-translational feedback circuit to limit the time course of expression of these stimulus-responsive transcription factors. Through a gain-of-function genetic screen, we identified a large group of potential targets of midnolin strongly enriched for nuclear proteins, especially transcriptional regulators, revealing that midnolin functions broadly to promote the degradation of proteins in the nucleus, where midnolin itself is predominantly localized.

The biological functions of midnolin are likely complex. Many of its substrates play central roles in the nervous and immune systems, and thus it will be important to establish the physiological function of midnolin in vivo. Midnolin was initially discovered because of its strong induction in the midbrain early during embryonic development (*39*). However, how midnolin expression is induced both during development and in cultured cells is currently unclear. Midnolin may have evolved to integrate various upstream stimuli and shape the proteome swiftly as a crucial response to the initial stimuli. Many midnolin substrates, including IEG proteins and lineage-specific transcription factors, undergo a transient burst of expression either in response to a particular stimulus or during a specific stage of development, and thus the prolonged expression of these transcription factors could be detrimental to organismal physiology. Indeed, previous studies have shown that chronic and repeated exposure to drugs of abuse such as cocaine leads to week-long accumulation of ΔFosB, which plays a crucial role for addiction in the nucleus accumbens of the brain (*35*–*37*). Our finding that ΔFosB is resistant to midnolin-dependent degradation provides a glimpse into the role that midnolin could play in brain function. Given that many IEG proteins are efficiently targeted for degradation by midnolin and that the precise expression of IEGs is critical for learning and memory (*1*, *3*, *40*), it is possible that disrupting or boosting midnolin function could affect the ability of animals to efficiently learn and store information in the brain.

IRF4, another midnolin substrate, is a lineage-specific transcription factor that is essential for the function and homeostasis of mature B and T cells and is an oncogenic driver of multiple myeloma (*15*, *16*, *41*). IRF4 protein expression is potently induced by diverse stimuli, including T cell receptor signaling and cytokines (*42*–*45*). How the IRF4 protein returns to baseline after stimulation was unknown, and our finding that IRF4 is potently targeted for degradation by midnolin may provide insights into the function of midnolin in the immune system. Taken together, these findings suggest that midnolin may serve as a key regulator that determines the precise expression kinetics of stimulus-induced transcription factors by controlling their protein stability in various cell types or tissues. In principle, midnolin could be used to target different proteins under different circumstances. For example, in flies, it was reported that a midnolin ortholog, Stuxnet, is cell cycle regulated and promotes Pc protein degradation in mitosis during development (*14*).

We found that midnolin promotes the degradation of its targets in a proteasome-dependent but ubiquitination-independent manner. This is supported by multiple lines of evidence. Removal of lysine residues from many midnolin targets did not abrogate the ability of midnolin to bind with and promote their destruction. A caveat to this interpretation is that ubiquitination can occur on the N-terminal amine of the first amino acid if it is not acetylated or, more rarely, on other residues such as cysteine, serine, or threonine (*24*). However, in contrast to the effects of proteasomal inhibition, blocking the E1 ubiquitin–activating enzyme did not abrogate midnolin function while stabilizing canonical substrates of the ubiquitin-proteasome system. These experiments allowed us to conclude that midnolin does not require ubiquitination for its degradative function. Instead, midnolin engages substrates using its Catch domain, which binds a hydrophobic region capable of β-strand formation that functions as a degron. Midnolin associates with the proteasome using its long, C-terminal α helix and contains an N-terminal ubiquitin-like domain that is essential to promote the degradation of bound substrates. We hypothesize a model whereby the concerted action of these three regions of midnolin allows for the ubiquitination-independent proteasomal degradation of midnolin-bound proteins (Fig. 7).

Previous work demonstrated that ornithine decarboxylase is targeted directly to the proteasome for degradation through a C-terminal unstructured sequence (*46*–*49*). However, the regions within substrates captured by midnolin are unlike any others previously described for E3 ubiquitin ligases. The midnolin degron appears to be generally unstructured but has the potential to form a β strand upon binding midnolin, with aliphatic residues of the degron buried within the hydrophobic core of the Catch domain. The integration of the β-strand degron into the Catch domain appears to complete a five-stranded antiparallel β-sheet structure. The somewhat degenerate nature of these hydrophobic degrons in various substrate proteins is likely explained by the fact that β sheets are stabilized not only by side chain interactions but also by backbone hydrogen bonding (*50*), which may reduce the need for specific amino acid interactions. A region with the propensity to form a β strand that is also biochemically compatible with the hydrophobic core of the Catch domain may be present in many proteins, thereby explaining how midnolin recognizes a diverse set of proteins. Given the ubiquity of such β strands as structural elements of proteins, midnolin recognition of unfolded proteins could behave as a general quality control mechanism.

How midnolin initiates the degradation of bound substrates is not completely understood mechanistically. We do not know if midnolin interacts with its targets before binding to the proteasome or if it associates with the proteasome constitutively and then recruits its substrates, thus defining a new subclass of proteosomes in the nucleus. We favor the latter possibility, because a unique feature of midnolin is that it interacts stably with the proteasome using a C-terminal α helix but not its Ubl domain. This is unlike the processivity factors Rad23 and ubiquilin, which bind ubiquitinated cargo and use their Ubl domain to interact with the proteasome (*29*, *30*, *51*–*53*). How the midnolin α helix binds the proteasome is currently unclear, and we do not yet understand how the Ubl functions. However, the Ubl domain of midnolin is necessary for the degradation of bound substrates, and it is noteworthy that midnolin itself is efficiently degraded by the proteasome in a ubiquitination-independent manner that requires the Ubl domain. Whether midnolin is reused or being degraded along with the substrate remains to be determined. We hypothesize that when midnolin interacts with a substrate, a conformational change occurs that allows the proteasome to recognize the Ubl

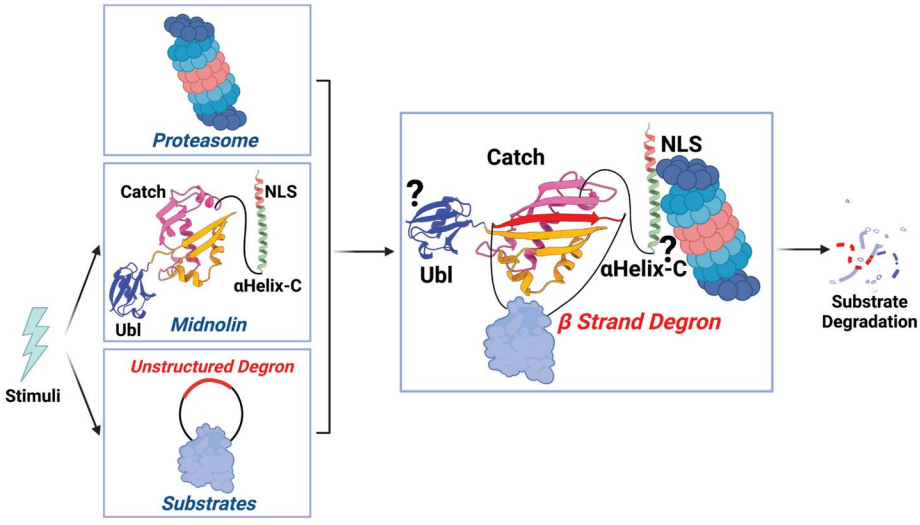


Fig. 7. Model for how midnolin functions to promote ubiquitination-independent proteasomal degradation. Midnolin is induced by growth factors and neurological stimuli, and its overexpression is sufficient to cause the degradation of its targets, including transcription factors such as c-Fos, FosB, ERG1, NR4A1, IRF4, and potentially many other proteins within the nucleus, where midnolin primarily resides. The degradation of its substrates does not require ubiquitination. Instead, midnolin uses its Catch domain to bind unstructured hydrophobic regions within substrates that have the potential to form a β strand that functions as a midnolin degron. Midnolin associates with the proteasome using its long C-terminal α helix and promotes the destruction of Catch-bound substrates through its N-terminal ubiquitin-like domain. Structures of the midnolin domains are derived from AlphaFold predictions. How the C-terminal α helix of midnolin binds the proteasome, whether a conformational change occurs after substrate binding, and how the ubiquitin-like domain confers degradative activity require further investigation (figure created with BioRender.com).

domain transiently to promote the degradation of the bound substrate with or without midnolin also being degraded. Structural and biochemical analyses are required to deepen our understanding of this noncanonical docking and degradative mechanism.

Our results suggest that the midnolin-proteasome pathway may represent a general mechanism by which the proteasome bypasses the traditional ubiquitination system to achieve selective degradation of many nuclear proteins. It has been reported that in bacteria, which do not contain the ubiquitin-proteasome system, a hierarchy of adaptors mediate selective degradation of diverse proteins by the proteasome-equivalent ClpXP protease complex (*54*, *55*). It will be of interest to determine whether, in the course of evolution, additional proteins have evolved to recruit proteins directly to the proteosome for degradation.

Materials and Methods

Cell culture

HEK-293T (ATCC, CRL-3216, RRID: CVCL_0063) and NIH/3T3 (ATCC, CRL-1658, RRID: CVCL_0594) cells were cultured at 37°C and 5% CO_2 in Dulbecco's modified Eagle's medium (DMEM) (Thermo Fisher Scientific, 11965118) supplemented with 100 units/ml penicillin, 0.1 mg/ml streptomycin (Thermo Fisher Scientific, 15070063), and 10% fetal bovine serum (Cytiva, SH30088.03). Ramos B cells (ATCC, CRL-1596, RRID: CVCL_0597) were cultured in RPMI 1640 medium (Thermo Fisher Scientific, A1049101) supplemented with 100 units/ml penicillin, 0.1 mg/ml streptomycin, and 10% fetal bovine serum. NIH/3T3 cells were starved overnight of serum and restimulated the following day with 20% serum for the indicated time points.

Animals were handled according to the protocol (IS00000074-6) approved by the Harvard University Office of the Institutional Animal Care and Use Committee, HMA Standing Committee on Animals and were in accordance with federal guidelines. Mouse cortical neurons were isolated and cultured as described previously (*56*). In short, embryonic cortices from wild-type C57BL/6NCrl mice (Charles River Laboratories strain number 027, five to 10 embryos, both males and females) were dissected at embryonic day 16.5 (E16.5) and dissociated with papain (Sigma Aldrich, 10108014001). After terminating the papain digestion with ovomucoid (trypsin inhibitor from Worthington), cells were triturated gently through a P1000 pipette before passing through a 40-μm filter, and then plated on cell culture dishes coated with poly-D-lysine (20 μg/ml) and laminin (4 μg/ml). The culture medium used for neurons was Neurobasal medium (Gibco) that contains 2% B27 supplement, penicillin-streptomycin (50 U/ml penicillin and 50 U/ml streptomycin), and glutaMAX (1 mM). The neurons were cultured at 37°C and 5% CO_2, treated with corresponding viruses at 3 days in vitro (DIV) while adding fresh culture medium at the same time (35% of total volume), silenced on 10 DIV overnight by the addition of 1 μM tetrodotoxin (TTX; Abcam, ab120055) and 100 μM AP5 (Thermo Fisher, 01-061-0), and harvested at 11 DIV after treatment with KCl stimulation buffer composed of 52.8 mM KCl, 0.62 mM $CaCl_2$, 0.31 mM $MgCl_2$, and 3.1 mM HEPES, pH 7.4, for the indicated times.

Cells were treated with 20 ng/ml PMA (Thermo Fisher Scientific, J-63916-MA), 10 μM MG132 (Selleckchem, S2619), or 500 nM TAK-243 (Selleckchem, S8341) from 1000× stock solution in dimethylsulfoxide (DMSO) for 6 hours unless stated otherwise.

Plasmids and cloning

The genome-wide CRISPR-Cas9 single-guide RNA (sgRNA) root library (five sgRNAs/gene, 94,335 sgRNAs total) was used previously, and the sgRNA information can be found in data S1 (*57*). The barcoded GPS ORFeome expression library was generated previously (*13*). The plasmids for human cDNAs containing a stop codon and conferring kanamycin resistance were obtained from the Ultimate ORF Clone collection (Thermo Fisher Scientific) in the form of Gateway entry clones: MIDN (IOH62653, BC094778.1), FOSB (IOH62162, NM_006732.2), FOS (IOH5624, NM_005252.3), ATF2 (IOH37849, NM_001880.2), CREB3 (IOH14437, BC009402.2), CREB5 (IOH53714, NM_001011666.1), ATF3 (IOH6465, NM_001674.2), BATF2 (IOH13295, BC012330.1), COMMD9 (IOH12792, NM_014186.3), C11ORF31 (IOH58679, NM_170746.2), ZNF621 (IOH44483, NM_198484.3), SERTAD2 (IOH42292, NM_014755.2), LMX1A (IOH34878, NM_177398.2), LMX1B (IOH34707, NM_002316.2), HOXD3 (IOH5660, NM_006898.4), CHCHD2 (IOH3869, NM_016139.2), FNDC3B (IOH10620, BC012204.1), SOX12 (IOH40697, NM_006943.2), TAX1BP3 (IOH13074, NM_014604.2), PAX4 (IOH34754, BC074761.2), FOXA3 (IOH10014, NM_004497.2), STYX (IOH10157, NM_145251.3), ZNF764 (IOH6451, NM_033410.2), RelB (IOH11686, NM_006509.2), PRRX1 (IOH36664, NM_006902.3), IRF1 (IOH2022, NM_002198.2), IRF2 (IOH10126, NM_002199.3), IRF8 (IOH42114, NM_002163.2), IRF9 (IOH28745, NM_006084.4), MYC (IOH2954, P01106), SPINDOC (IOH28799, NM_138471.1), PPDPF (IOH4080, NM_024299.2), PAX8 (IOH3823, NM_003466.3), FOXS1 (IOH13387, NM_004118.3), NEUROD1 (IOH3394, NM_002500.2), MIER2 (IOH40210, NM_017550.1), IRF4 (IOH12141, NM_002460.2), GATA1 (IOH57792, NM_002049.3), CITED1 (IOH5542, BC004240.1), and XRCC1 (IOH40644, NM_006297.2).

The plasmids for human cDNAs lacking a stop codon and conferring spectinomycin resistance were obtained from the Human ORFeome library V8.1 (Dana Farber Cancer Institute) in the form of Gateway entry clones: EGR1 (ORF_ID no. 14665, BC073983.1) and NR4A1 (ORF_ID no. 292, BC016147.1).

CBX4 and hemagglutinin (HA)–tagged wild-type or K-to-R mutated cDNA were generated by synthesis using Integrated DNA Technologies (IDT) with attB1 and attB2 overhangs for cloning into the pDONR221 (Thermo Fisher Scientific, 12536017) through a BP recombination reaction (Thermo Fisher Scientific, 11789020) to generate the entry clone. Similarly, the following attB1 and attB2 overhangs we included in primers to generate fragments for sufficiency experiments by polymerase chain reaction (PCR) for cloning into the pDONR221: attB1, GGGGACAAGTTTGTACAAAAAAGCAGGCTTAgccacc; attB2, GGGGACCACTTTGTACA-AGAAAGCTGGGTA.

Entry clones were mutagenized by PCR using the Q5 Site Directed Mutagenesis kit (NEB, E0554S) and the primers for mutagenesis were designed using the NEBaseChanger program.

The midnolin amino acid sequence used was as follows: MEPQPGGARSCRRGAPGGACELGPAAEAAPMSLAIHSTTGTRYDLAVPPDETVEGLRKRLSQRLKVPKERLALLHKDTRLSSGKLQEFGVGDGSKLTLVPTVEAGLMSQASRPEQSVMQALESLTETQVSDFLSGRSPLTLALRVGDHMMFVQLQLAAQHAPLQHRHVLAAAAAAAAAARGDPSIASPVSSPCRPVSSAARVPPVPTSPSPASPSPITAGSFRSHAASTTCPEQMDCSPTASSSASPGASTTSTPGASPAPRSRKPGAVIESFVNHAPGVFSGTFSGTLHPNCQDSSGRPRRDIGTILQILNDLLSATRHYQGMPPSLAQLRCHAQCSPASPAPDLAPRTTSCEKLTAAPSASLLQGQSQIRMCKPPGDRLRQTENRATRCKVERLQLLLQQKRLRRKARRDARGPYHWSPSRKAGRSDSSSSGGGGSPSEASGLGLDFEDSVWKPEANPDIKSEFVVA

Midnolin regions for truncations or sufficiency experiments were as follows: ΔUbl, residues 31 to 105; ΔCatch1, residues 112 to 156; ΔCatch2, residues 266 to 332; ΔNLS, residues 402 to 413; ΔαHelix-C, residues 377 to 413; αHelix-C region fused to MBP, residues 360 to 432; Catch domain alone for sufficiency experiment, residues 102 to 334; ΔLoop1, residues 160 to 260; ΔLoop2, residues 163 to 257; ΔLoop3, residues 169 to 251; Catch1 fused to 2xHA-GFP, residues 107 to 166; Catch2 fused to 2xFLAG-MBP, residues 254 to 337; FosB truncations: deletion1, residues 2 to 42; deletion2, residues 43 to 82; deletion3, residues 83 to 122; deletion4, residues 123 to 149; deletion5, residues 150 to 237; ΔFosB, residues 238 to 338; predicted β strand truncations: FOXS1, residues 245 to 260; CBX4, residues 538 to 558; NEUROD1, residues 277 to 288; SPINDOC, residues 314 to 328; IRF4, residues 208 to 229; FOSB (N-terminal β strand), residues 67 to 75; EGR1, residues 128 to 145; β-strand sufficiency peptides: EGR1, residues 113 to 172; IRF4, residues 192 to 248; FosB, residues 238 to 338.

Entry clones were subcloned into the following lentiviral Gateway destination vectors using an LR recombination reaction (Thermo Fisher Scientific, 11791100): pHAGE-GPS 3.0 (*13*), pHAGE-GPS 3.2 (*58*), pHAGE-CMV-2xFLAG-N (this paper), pHAGE-CMV-2xHA-N (this paper), pHAGE-CMV Puromycin (*59*), pHAGE EF1α BFP (this paper), CMV-C-2xFlag expression vector (Addgene, 118372), pHAGE TRE Blasticidin (this paper), pHAGE-CMV-2xHA-GFP-N (this paper), pHAGE-CMV-2xFLAG-GFP-N (this paper), and pHAGE-CMV-2xFLAG-MBP-N (this paper). The pInducer20 mouseCD19 (*59*) plasmid was used to generate dox-on responsive HEK-293T cells.

Lentiviral CRISPR-Cas9 vectors containing the sgRNA of interest were cloned by first digesting the backbone lentiCRISPRv2 (Addgene, 52961) using BsmBI (NEB, R0739S). Then, sgRNA oligos containing CACC or AAAC overhangs were obtained from IDT, phosphorylated and annealed together, and ligated into the linear backbone by T4 (NEB, M0202S) ligation. Sequences were as follows: nontargeting sgControl, GTATTACTGATATTGGTGGG; human sgMIDN #1, GAAGCTGCAGGAGTTCGGCG; human sgMIDN #2, GCTGACCTTGGTACCCACCG; and mouse sgMIDN, GCGAGCTGAACACGGCCA.

Lentivirus production

Lentivirus was generated by transfecting HEK-293T cells using PolyJet (SignaGen, SL100688) following the manufacturer's instructions with plasmids encoding Tat, Rev, Gag-Pol, and VSV-G and lentiviral transfer vectors. Specifically, plasmid DNA was diluted into DMEM lacking supplements, and 3 μl of PolyJet reagent was used per 1 μg of plasmid DNA. One day after transfection, the medium was removed and the cells were supplemented with fresh complete culture medium. The lentiviral supernatant was collected 48 and 72 hours after transfection, passed through a 0.45-μm filter, and either applied directly to cells or stored at −80°C for later use. For most experiments, lentivirus was packaged in six-well plates. For library preparations, lentivirus was packaged in eight 15-cm plates using 13 μg of total DNA per plate for transfection, pooled, concentrated using the lenti-X concentrator (Takara, 631232), and aliquoted.

Midnolin overexpression and flow cytometry

Cells were seeded in six-well plates at 200,000 cells/well and transfected 2 days later using Polyjet with EF1α midnolin coexpressing BFP or BFP alone as a negative control. The medium was changed 1 day after transfection, and cells were analyzed 2 days after transfection.

Cells were prepared for flow cytometry by aspirating old media and rinsing cells once with phosphate-buffered saline (PBS). The cells were detached using 0.05% trypsin at room temperature, and the trypsin was neutralized using fresh complete media. The cells were then analyzed on a CytoFLEX S flow cytometer (Beckman Coulter, V2-B2-Y3-R2 version C09762) and the CytExpert software (Beckman Coulter) was used to collect flow cytometry data. All flow cytometry data were analyzed using FlowJo software. For FACS, a Sony MA900 was used for routine sorting of single clones, and a MoFlo Astrios (Beckman Coulter) instrument was used to collect cells for the CRISPR-Cas9 and GPS ORFeome screens.

Generating doxycycline-inducible HEK-293T cells

Wild-type and *MIDN* KO HEK-293T cells were infected with lentivirus encoding the pInducer20 system expressing mouse CD19 as a doxycycline-inducible marker for cell surface staining. After infection, the cells were treated with doxycycline (100 ng/ml) for 2 days and then stained on the cell surface using phycoerythrin (PE) antimouse CD19 antibody (BioLegend, 152407, RRID: AB_2629816) for 30 min using 1 μl of antibody diluted in 100 μl of medium containing 1 million cells. The cells were rinsed twice with PBS before sorting for the PE-positive population. The sorted cells were allowed to expand for >1 week before being restained using PE antimouse CD19 antibody in the absence of doxycycline induction. The PE-negative population was sorted, and the cells were expanded before another round of staining and sorting for the PE-negative population. This allowed for a population of cells that responded to doxycycline with minimal leakiness. Finally, lentivirus encoding the pHAGE TRE-midnolin was infected into these *MIDN* KO HEK-293T cells and selected with blasticidin to allow for stable, doxycycline-inducible expression of midnolin.

EGR1 and FosB CRISPR-Cas9 screens

Genome-wide CRISPR-Cas9 screens were performed to uncover regulators of EGR1 and FosB protein stability. Specifically, the plasmid library was packaged into lentivirus by transfecting HEK-293T cells using PolyJet as described earlier, and the lentivirus was titered to obtain a multiplicity of infection ~0.3. HEK-293T cells were generated to express the GPS 3.0 FosB or GPS 3.2 EGR1 reporters by selecting using hygromycin (200 μg/ml). These cells were then transduced with the titered CRISPR-Cas9 genome-wide Root library lentivirus at a multiplicity of infection (MOI) of ~0.3 to maintain a 500× representation throughout. Cells were selected 48 hours after transduction for 7 days using puromycin (2 μg/ml) to remove uninfected cells. On the ninth day of puromycin selection, the 95th percentile most stable cell population was collected based on the GFP/DsRed ratio by FACS using a MoFlo Astrios instrument (Beckman Coulter). Additionally, the unsorted input cells were collected based on the number of cells collected in the enriched population. Collected cells were rinsed once with PBS, pelleted, and stored at −80°C.

Midnolin GPS ORFeome screen

The GPS ORFeome screen was performed as described previously (*13*) with some modifications. Sufficient cell numbers were used to

maintain at least a 300-fold coverage of the library throughout. The library was packaged into lentivirus and used to transduce *MIDN* KO HEK-293T at a MOI of 0.2. Two days after transduction, the HEK-293T cells were treated with 2 μg/ml puromycin for 6 days to remove uninfected cells, passaging the library once in between the selection period. The library-expressing cells were plated at 4 million cells/plate in a 15-cm dish and transfected 2 days later using Polyjet with 8 μg of DNA from EF1α-midnolin coexpressing BFP or BFP alone as a negative control. The cells were harvested 2 days after transfection and sorted into six stability bins based on the GFP/DsRed ratio by FACS using a MoFlo Astrios instrument (Beckman Coulter). The sorting gates were established using the BFP control to ensure that one-sixth of the population was collected per bin. Once the control populations were collected, the cells overexpressing midnolin were partitioned using the exact same sorting and gating settings as the control. The collected cells from each stability bin were rinsed once with PBS, pelleted, and frozen at −80°C for at least 12 hours.

Deconvolution of the pooled screens

Cell pellets were thawed and genomic DNA was harvested using a Gentra Puregene Core Kit (Qiagen). The sgRNAs or barcodes were then amplified by PCR using all the genomic DNA as a template (4 μg of DNA per reaction) to include stagger sequences and Q5 Hot Start High-Fidelity DNA Polymerase from NEB. A second round of PCR was performed using the clean PCR1 product to add the Illumina P5 and P7 adapter sequences. PCR2 samples were cleaned, pooled in the correct ratio, and sequenced on a NextSeq 500 instrument. The abundance of sgRNAs or barcodes were extracted from the raw sequencing data using Cutadapt (*60*) and mapped onto the reference library using Bowtie2 (*61*).

MAGeCK was used to determine the enrichment of sgRNAs in the 95the percentile relative to the input population (*62*). The MAGeCK score plotted on the *y*-axis (Fig. 1, B and C) represents the negative log10 of the "pos|score" value generated by MAGeCK.

For the GPS ORFeome analysis, the abundance of each ORF was corrected to account for sequencing depth and a protein stability index (PSI) score between 1 (most unstable) and 6 (most stable) was calculated using the following formula for each extracted ORF:

$$\mathrm{PSI} = \sum_{i=1}^{6} \mathrm{R}i * i$$

where i is the number of the stability bin denoted as an integer and Ri is the Illumina read proportion extracted from the bin i. The change in protein stability between midnolin and BFP is denoted as the difference in PSI (ΔPSI).

Gene set enrichment of GPS ORFeome hits

Gene set enrichment analysis (GSEA) was performed with GSEAPreranked (*63*) (v4.3.2) using rank weights derived from GPS ORFeome ΔPSI values as input. These were tested for enrichment across the Human Molecular Signatures Database (*64*) (MSigDB v2022.1.Hs) C5 Gene Ontology collection (i.e., GO:BP, GO:MF, and GO:CC). The classic Kolmorogorov-Smirnov scoring scheme was used with 10,000 permutations and excluded gene sets with <10 or >1000 entries when intersected with the list of GPS ORFs.

Generating *MIDN* KO cells

To generate isogenic single clones lacking midnolin, HEK-293T cells were transfected with the lentiCRISPRv2 BFP plasmid encoding the given sgRNA of interest using Polyjet. Several days after transfection, the BFP-positive cells were collected as single cells into 96-well plates using FACS. The single clones were allowed to expand for 2 weeks before screening for a KO phenotype by immunoblotting and next-generation sequencing of the genomic DNA locus encompassing the cut site (fig. S11A).

To generate a population-level depletion of midnolin in mouse NIH/3T3 fibroblasts, cells were transduced with lentivirus encoding the given mouse lentiCRISPR v2 sgRNA coexpressing puromycin. Three days after transduction, cells were selected by 2 μg/ml puromycin for 5 days and expanded for further analysis.

To generate a population-level depletion of midnolin in primary cortical neurons, cells were transduced with lentivirus encoding the given mouse lentiCRISPRv2 sgRNA coexpressing puromycin.

To generate a population-level depletion of midnolin in Ramos B cells, the cells were transduced by spinfection with lentivirus encoding the given lentiCRISPRv2 sgRNA#1 coexpressing BFP. Specifically, the cells were incubated with lentivirus for 30 min with centrifugation at 2000 rpm at room temperature. The cells were then allowed to expand for 6 days after transduction, and the BFP-positive cells were collected by FACS.

Generating 3xHA-MIDN knock-in HEK-293T cells

To generate an endogenous 3xHA-tagged midnolin cell line, we used homology-directed repair and CRISPR-Cas9. We reasoned that an N-terminal epitope tag would be tolerable, because an N-terminal tagged midnolin transgene could interact with substrates and promote their degradation. To establish the knock-in line, the following custom Alt-R sgRNA and single-stranded DNA (ssDNA) homology-directed templates were generated by synthesis from IDT. The sgRNA template was: rCrCrGrGrGrCrUrGrCrGrGrCrUrCrCrArUrCrCrGrUrUrUrArArGrArGrCrUrArUrGrCrUrGrGrArArArCrArGrCrArUrArGrCrArArGrUrUrUrArArArUrArArGrGrCrUrArGrUrCrCrGrUrUrArUrCrArArArArArGrUrGrGrCrArCrCrGrArGrUrCrGrGrUrUrUrUrUrU. The ssDNA template was: CGGCGCCCGCCGCCCCCAGCCCCCCAGCGCGCGCCGGGGATGTATCCCTATGACGTGCCTGATTACGCCGGCGGAGGATCCTACCCCTATGATGTGCCTGACTACGCTGGCAGCGGAGGATACCCTTATGATGTGCCTGATTATGCTGGAGGTGGAGGTAGTGAGCCGCAGCCCGGCGGCGCCCGGAGCTGCCGGCGCGG.

spCas9, sgRNA, and ssDNA template were introduced to cells by nucleofection. Specifically, per nucleofection of 100,000 cells in Lonza strip nucleofector system (V4XC-2032), 0.8 μl of 62.1 μM spCas9 (Aldevron, 9212-0.25MG), 0.8 μl of 100 μM sgRNA (Alt-R from IDT), 0.25 μl of 10× buffer 3.1 (NEB, B6003S), and 0.65 μl of H_2O were mixed to bring the final volume to 2.5 μl per nucleofection and incubated Ribonucleoprotein (RNP) at room temperature for 30 min. Making a master mix to troubleshoot the ssDNA template concentration is recommended to avoid small volumes. Then, 2 μl of the RNP was added to 100,000 HEK-293T cells resuspended in 16.4 μl of Nucleofector solution plus 3.6 μl of supplement. Donor DNA was added directly to this solution at a final concentration of 500 nM or 2 μM. Cells were nucleofected using a 4D Nucleofector X Unit, and a GFP-positive control was included to ensure that the nucleofection worked properly. Once nucleofected, the cells were allowed to recover for 10 min at room temperature before being added to six-well plates containing warm medium. The cells were allowed to expand for several days, and a cell lysate was collected to ensure the endogenous editing worked by immunoblotting. Then, single cells were partitioned into 96-well plates to obtain isogenic clones, which were validated by immunoblotting once expanded to ensure successful HA knock-in.

Immunoprecipitation

Cells stably expressing the indicated epitope-tagged protein were cultured in 10- or 15-cm plates and allowed to reach 90% confluency before lysis. Alternatively, cells in 10-cm dishes were transiently transfected with 3 μg of the indicated plasmid DNA using Polyjet when 50% confluent. The medium was changed 1 day after transfection, and the cells were lysed 2 days after transfection with or without any necessary perturbations such as proteasomal inhibition. For lysis, cells were rinsed once with ice-cold PBS by pouring and collected by scraping in 0.7 ml (10 cm plate) or 1 ml (15 cm plate) of lysis buffer containing 0.5% CHAPS, 40 mM HEPES, pH 7.4, 100 mM NaCl, and 4 mM EDTA supplemented with 1× protease and a phosphatase inhibitor cocktail (Thermo Fisher Scientific, 78441). Cell lysates were incubated with end-to-end rotation at 4°C for 30 min before clarification by centrifugation at 21,000*g* for 15 min at 4°C. Anti-FLAG (Sigma, M8823,

RRID: AB_2637089) or anti-HA (Thermo Fisher Scientific, 88836, RRID: AB_2749815) magnetic beads were rinsed three times in lysis buffer using 15 μl of beads for every harvested plate. A 50-μl aliquot of the cell lysate was collected as input, and the remaining supernatant was incubated with the beads for 2 hours at 4°C with end-to-end rotation. The immunoprecipitants were washed three times with the same lysis buffer, and the cell lysates and immunoprecipitants were resuspended in Tris-glycine SDS sample buffer (Thermo Fisher Scientific, LC2676) containing 10% 2-mercaptoethanol. Protein was eluted by heating at 95°C for 4 min before analysis of protein content by immunoblotting.

Immunoblotting

For immunoprecipitation experiments, the collected samples were loaded directly for immunoblotting. For measuring steady-state abundance changes, cultured cells were lysed using 1× RIPA buffer (Boston BioProducts, BP-115X) supplemented with 1× protease and a phosphatase inhibitor cocktail (Thermo Fisher Scientific, 78441) for 15 min at 4°C. Samples were centrifuged at 21,000*g* for 15 min at 4°C, and the protein concentration was normalized using a bicinchoninic acid assay (Thermo Fisher Scientific, 23225). Clarified supernatants were resuspended in Tris-glycine SDS sample buffer (Thermo Fisher Scientific, LC2676) containing 10% 2-mercaptoethanol. Samples were loaded into 4 to 12% Tris-glycine 15-well precast gels (Thermo Fisher Scientific, XP04125BOX), and electrophoresis was run in 1× Tris-glycine SDS running buffer (Thermo Fisher Scientific, LC2675-4) at a constant 165 to 180 volts until the molecular weight ladder (Thermo Fisher Scientific, 26619) ran to the bottom. The protein within the gel was transferred to a 0.2-μm nitrocellulose membrane (Bio-Rad, 170-4158) using the Trans-Blot Turbo Transfer System (Bio-Rad). Nitrocellulose membranes were then blocked using 5% milk (LabScientific, M-0842) diluted in 1× Tris-buffered saline plus Tween 20 (TBST) (CST, 9997S) for at least 30 min at room temperature with gentle rocking. Primary antibodies were then diluted directly in the blocking solution at a 1:1000 dilution and incubated overnight at 4°C with gentle rocking. The following primary antibodies were used: rabbit anti-EGR1 (CST, 4153, RRID: AB_2097038), rabbit anti-FosB (CST, 2251, RRID: AB_2106903), rabbit anti-c-Fos (in house) (*65*), rabbit anti-NR4A1 (in house, warning has high background), rabbit anti-Midnolin (Proteintech, 18939-1-AP, RRID: AB_2878569), rabbit anti-PSMD2 (CST, 25430, RRID: AB_2798903), rabbit anti-PSMA2 (CST, 2455, RRID: AB_2171400), rabbit anti-HA (CST, 3724, RRID: AB_1549585), rabbit anti-FLAG (CST, 14793, RRID: AB_2572291), rabbit anti-mTOR (CST, 2983, RRID: AB_2105622), rabbit anti-Actin (CST, 4970, RRID: AB_2223172), rabbit anti-GAPDH (CST, 5174, RRID: AB_10622025), rabbit anti-CBX4 (CST, 30559, RRID: AB_2798991), rabbit anti-CBX8 (CST, 14696, RRID: AB_2687589), rabbit anti-ATF2 (CST, 35031, RRID: AB_2799069), rabbit anti-ATF3 (CST, 33593, RRID: AB_2799039), rabbit anti-IRF1 (CST, 8478, RRID: AB_10949108), rabbit anti-RelB (CST, 4922, RRID: AB_2179173), rabbit anti-STAT3 (CST, 9139, RRID: AB_331757), rabbit anti-SPINDOC (Sigma, HPA040128, RRID: AB_10673027), rabbit anti-XRCC1 (CST, 2735, RRID: AB_2218471), rabbit anti-CITED1 (Proteintech, 26999-1-AP, RRID: AB_2880718), rabbit anti-SOX12 (Proteintech, 23939-1-AP, RRID: AB_2879368), rabbit anti-FOXP3 (CST, 5298, RRID: AB_10839127), rabbit anti-c-Myc (CST, 5605, RRID: AB_1903938), rabbit anti-p27 (CST, 3686, RRID: AB_2077850), rabbit anti-Ubiquitin (CST, 43124, RRID: AB_2799235), and rabbit anti-IRF4 (CST, 4299, RRID: AB_10547141).

After overnight incubation, the blots were rinsed four times quickly and three additional times for longer, 10-min incubations using 1× TBST. After rinsing, the blots were incubated with 5% milk in 1× TBST and the following secondary antibodies were applied directly at a 1:2000 dilution: anti-rabbit IgG, HRP-linked (CST, 7074, RRID: AB_2099233) or anti-mouse IgG, HRP-linked (CST, 7076, RRID: AB_330924). The blots were incubated in secondary antibody for 1 hour at room temperature with gentle rocking before rinsing as done for the primary antibody. The blots were exposed to either Pierce ECL Western blotting substrate (Thermo Fisher Scientific, 32106) for strong antibodies or highly abundant protein, or Immobilon western chemiluminescent HRP substrate (Sigma, WBKLS0500) for weaker antibodies or less abundant proteins. All immunoblotting data were collected using high-sensitivity autoradiography film (Denville Scientific, E3218).

We note that the midnolin antibody (Proteintech, 18939-1-AP) has limitations that are important for readers to consider. The protein levels of endogenous midnolin appear quite low in most cell types, and this antibody contains too much background (nonspecific bands) to robustly detect endogenous midnolin protein from cell lysates at steady state. However, endogenous midnolin protein can be detected with this antibody if the cells are pretreated for a few hours with 10 μM MG132 or if the cells are overexpressing midnolin (fig. S11B).

Mass spectrometry of endogenous midnolin immunoprecipitants

HEK-293T cells expressing endogenous 3xHA-tagged midnolin were cultured to 90% confluency in five 15-cm plates per condition, and unedited wild-type HEK-293T cells were cultured in five 15-cm plates. The knock-in cells were treated was DMSO or 10 μM MG132 for 6 hours, with the unedited wild-type HEK-293T serving as the background, and treated with 10 μM MG132 for 6 hours. An anti-HA immunoprecipitation was performed using the same lysis conditions and protocol as described in the immunoprecipitation section. After the final wash, the beads were resuspended in 100 μl of 50 mM Tris, pH 8.5, containing 5% SDS, and the samples were heated at 95°C for 5 min to elute the proteins.

Eluted proteins were then digested using trypsin on S-Trap Micro columns (Protifi, C02-micro-10) following the manufacturer's protocol. Specifically, proteins were first reduced using 5 mM Tris(2-carboxyethyl)phosphine for 15 min at 55°C, and then alkylated with 20 mM iodoacetamide for 30 min in the dark at room temperature. After alkylation, the samples were acidified using phosphoric acid to a final concentration of 2.5% (v/v) and 10 volumes of 100 mM Tris, pH 7.55 in 90% methanol/10% water were added to the samples to dilute the protein. This solution was then passed through S-Trap column by centrifuging for 30 s at 4000*g*. Multiple rounds of centrifugation were needed to load the entirety of one sample onto one column. Once the protein was trapped, the column was rinsed three times using 100 mM Tris, pH 7.55, in 90% methanol and 10% water, followed by a dry spin, before adding 2 μg of trypsin suspended in 20 μl of 50 mM ammonium bicarbonate, pH 8. Columns were kept overnight at 37°C in a humid environment. After digestion, the peptides on the column were eluted by centrifuging three times for 1 min at 4000*g* using three buffers applied sequentially: (i) 40 μl of ammonium bicarbonate, pH 8; (ii) 40 μl of 0.2% formic acid in water; and (iii) 40 μl of 50% acetonitrile in water. The pooled peptides were dried under reduced pressure using a SpeedVac and resuspended in 30 μl of 0.1% formic acid in water. Liquid chromatography–tandem mass spectrometry data were acquired as reported previously (*66*) by injecting 10 μl of resuspended peptide sample.

A protein database consisting of the Human UniProt SwissProt proteome (downloaded on 13 November 2022) was used to identify proteins that coimmunoprecipitated with endogenous 3xHA-midnolin. Specifically, the FragPipe graphical user interface (v18.0) was used to search the data using the MSFragger search engine and to perform postprocessing of the search results. The following parameters were used in the search. Tryptic peptides with a maximum of two missed cleavages were considered. Additionally, carbamidomethylation of cysteine was set as a fixed modification, and oxidation of methionine was allowed as a variable modification, with a maximum of four variable modifications per peptide. The allowed mass tolerances were 10 ppm for precursor ions and 0.04 Da for product ions. Peptide hits were filtered to a false discovery rate of 1% using PeptideProphet as implemented in FragPipe.

Immunofluorescence

For experiments in fig. S5, A and B, 400,000 HEK-293T cells with indicated genetic background

were plated on poly-D-lysine–coated coverslips (Ted Pella, Inc.). On the following day, indicated treatments with DMSO, MG132, or PMA were performed for 6 hours before collection. Culture medium was aspirated and cells were washed with PBS once before fixation by 4% paraformaldehyde in PBS for 15 min at room temperature. After three PBS washes, cells were permeabilized with 0.05% Triton X-100 in PBS for 5 min at room temperature. Cells were washed with PBS three times and placed in immunofluorescence blocking buffer (LI-COR, 927-70001) for 45 min at room temperature. Primary antibodies mouse anti-FLAG (Sigma, F1804, RRID: AB_262044) and rabbit anti-HA (CST, 3724, RRID: AB_1549585) were diluted 1:400 in the blocking buffer and added on top of coverslips with cells for an overnight incubation at 4°C. After being rinsed three times with PBS, the cells were incubated in the dark at room temperature with secondary antibodies (1:500 dilution) and 8 μM Hoechst 33342 dye (Thermo Fisher Scientific, H3570), both diluted in blocking buffer, for 1 hour. Coverslips were rinsed thoroughly using PBS and mounted on glass slides using ProLong gold antifade mountant (Thermo Fisher Scientific, P10144).

An Alexa Fluor 488–conjugated secondary antibody was used for the FLAG (Thermo Fisher Scientific, A-11001, RRID: AB_2534069) and HA (Thermo Fisher Scientific, A-11008, RRID: AB_143165) staining in fig. S5, A and B, and the excitation wavelength was 488 nm. The excitation wavelength of the Hoechst nucleus-staining dye was 405 nm.

Image acquisition was done by a Zeiss AxioVert200M microscope with a 100× oil-immersion objective, a Yokogawa CSU-22 spinning-disk confocal head with a Borealis modification (Spectral Applied Research/Andor), and a Hamamatsu ORCA-ER CCD camera. Image acquisition and hardware were controlled by the MetaMorph software package (Molecular Devices). The excitation lasers used to capture the images were 405 and 488 nm.

RNA extraction and quantitative PCR

Total RNA was extracted from cells with the RNeasy Plus Mini Kit (Qiagen, 74134), and cDNA was generated from freshly extracted RNA using the iScript cDNA Synthesis Kit (Bio-Rad, 1708891) following the manufacturer instructions for both kits. Specifically, 250 ng of RNA was used for a 20 μl reaction to generate the cDNA. Platinum SYBR Green quantitative PCR (qPCR) Supermix-UDG (Thermo Fisher Scientific, 11733038) and 2 μl of cDNA was used for qPCR reactions. Specifically, master mixes were prepared to contain 10 μl of SYBR, 7.5 μl of water, and 0.5 μl of 40× primers per 20 μl of qPCR reaction. Quantstudio 6 Pro (Thermo Fisher Scientific) was used to run the qPCR reactions. The following intercalating premixed qPCR primers were obtained from IDT. Mouse MIDN (Mm.PT.58.10544931): GCGTCAACTTGCTCCCAT, AACGCCTCAAAGTACCCAAG; mouse EGR1 (Mm.PT.58.29064929): GATAACTCGTCTCCACCATCG, AGCGCCTTCAATCCTCAAG; mouse c-Fos (Mm.PT.58.29977214): GGCACTAGAGACGGACAGAT, ACAGCCTTTCCTACTACCATTC; mouse FosB (Mm.PT.58.10990878): AGAGACACTTACCCCAGAAGA, GCTCTGCCTTTTCCTCTTCA; mouse Actin (Mm.PT.39a.22214843.g): GACTCATCGTACTCCTGCTTG, GATTACTGCTCTGGCTCCTAG; mouse mTOR (Mm.PT.58.28403918): TGCATCACTCGTTCATCCTG, AAGTCATCACATCCAAGCAGA.

Changes in mRNA levels were determined by subtracting the cycle quantification (Cq) values generated during the qPCR between the gene of interest and the control to yield a ΔCq value. Data were then normalized to the indicated control condition by substracting the ΔCq values by the average ΔCq of the indicated control condition to generate the ΔΔCq. Plotted in graphs are $2^{-\Delta\Delta Cq}$ from three biological replicates and the following statistical tests were used.

For Fig. 2, E and F, data were analyzed using an ordinary one-way ANOVA followed by a Tukey's multiple-comparisons test where $^{****}P < 0.0001$. For fig. S1D, data were analyzed using a two-way ANOVA followed by Šidák's multiple comparisons test, where ns is not significant and $^{***}P < 0.001$.

Computational identification of substrate β strands

AlphaFold multimer predictions

To identify β strands within hits identified in the ORFeome GPS screen, genes with ΔPSI < −0.5 were taken (n = 508) and the longest sequence across corresponding protein accession IDs (either NCBI Reference Sequence or Ensembl ID) was used as the input sequence for downstream steps (because barcodes from the screen were grouped at the gene level but could represent multiple isoforms). These sequences were individually paired with the MIDN sequence (UniProtKB: Q504T8) as a two-sequence FASTA file input into AlphaFold (v2.2.0) for multimer prediction with default reference databases specified as in (*26*) and max_template_date=2022-01-01. Any selenocysteines were recoded as cysteines, and three substrates (ACSBG2, ACSS2, and RIMBP3) that failed MSA using the default settings were rerun successfully by replacing the UniClust30_2018_08 database with UniRef30_2022_02.

Identification of substrate β strands within midnolin β sheet

The 25 ranked PDB models from each AlphaFold run with MIDN and one of the substrates were then processed by a custom Python script to identify PDB models that folded a linear stretch of the substrate into β-strand conformation placed between β strands of the corresponding MIDN domain. In more detail, a pairwise distance matrix was first computed between each α-carbon atom in MIDN and each α-carbon atom in the substrate as

$$D_{i,j} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2}$$

where x_i, y_i, and z_i are the coordinates of the i^{th} substrate α-carbon atom and x_j, y_j, and z_j are the coordinates of the j^{th} MIDN α-carbon atom. Because most β sheets have interstrand distances <5 Å (*67*), the distance matrix was scanned to identify sequential substrate residues <5.5 Å from corresponding linear stretches within each adjacent MIDN β strand (ie. $D_{i,j} <$ 5.5 for both some sequential set of i with some sequential set of j, where 148 ≤ j ≤157, as well as the same set of i with another sequential set of j, where 279 ≤ j ≤ 286).

Secondary structure assignment for the PDB model was done with the DSSP algorithm (*68*). Substrate residues satisfying the distance requirements specified above were then retained if they were assigned the extended β strand secondary structure (i.e., "E" coding). Because DSSP relies on flanking residues to call secondary structure, the most N- and C-terminal residues are not assigned secondary structure. To avoid excluding them from β-strand assignments, they were assigned "E" coding if the adjacent residue had been assigned "E" coding. To catch residues that are part of a β strand but slightly further from one or both of the MIDN β strands, this set of residues was then expanded by seven residues in each direction, and, again, only those with an extended β-strand secondary structure were kept. Finally, the longest contiguous stretch of β-strand secondary structure was kept (if any) for final reporting (data S4).

Properties of identified substrate β strands

Relative disorder comparison

ORFs used in AlphaFold multimer folding alongside midnolin were matched with UniProtKB accessions using UniParc to find identical proteins that had already been folded as monomers in the AlphaFold Protein Structure Database (https://alphafold.ebi.ac.uk/). This yielded existing models for 126 of 205 substrates predicted to interact with the midnolin Catch domain, for which the predicted local distance difference test (pLDDT) scores were extracted from corresponding PDB files. The pLDDT scores for these residues interacting with the midnolin Catch domain were averaged and compared with the average of the rest of the substrate (paired t test, $P = 2.72 \times 10^{-9}$) as an approximation of disorder in the original substrate (*69*), where lower pLDDT scores correspond to increased disorder. The regions that are predicted to interact with the Catch domain are predicted to be unstructured by AlphaFold in the native, midnolin-free state.

Amino acid enrichment

For each residue, n, going into the β strand from either the N- or C-terminal side, the overall frequency of each amino acid for β strands of length ≥2n (to avoid double counting) was normalized by the background amino acid frequency across substrate sequences. Frequencies were then computed for amino acids before and after all β strands provided the β strand was not the N or C terminus of the protein, respectively.

Hydrophobicity

The relative hydrophobicity of residues comprising the MIDN-interacting β strand was assessed by a two-sided paired *t* test between the mean hydrophobicity index (*70*, *71*) (at pH 7) of residues in the β strand with that of residues comprising the rest of the substrate.

Identification of MIDN-facing β-strand side chains

Because β strands within a β sheet make contacts with each other through backbone interactions, where side chains alternatingly project above and below the plane of the β sheet, residues composing the MIDN-interacting β strand can be parsed into those facing toward or away from the MIDN Catch domain. Side chains of the substrate β strand facing the MIDN Catch domain were identified by first generating another distance matrix as before, but between each α-carbon atom in MIDN and each β-carbon atom in the substrate. For each residue in the identified substrate β strand, those with β-carbon distance to MIDN Catch domain (approximated by α-carbon position for MIDN isoleucine residue 309) less than their α-carbon distance were annotated as MIDN-facing. Because glycine lacks a β-carbon and residues at the ends of β strands may have more rotational variability, but side chain orientations along the β strand should alternate between facing toward or away from the MIDN Catch domain, a filter was then applied to determine whether the identified MIDN-facing side chains better matched either the set of odd or set of even residues. MIDN-facing side chains were then annotated as that set. Properties such as relative frequency and hydrophobicity were then computed for inward- and outward-facing side chains.

REFERENCES AND NOTES

1. K. Minatohara, M. Akiyoshi, H. Okuno, Role of immediate-early genes in synaptic plasticity and neuronal ensembles underlying the memory trace. *Front. Mol. Neurosci.* **8**, 78 (2016). pmid: 26778955
2. S. B. McMahon, J. G. Monroe, The role of early growth response gene 1 (egr-1) in regulation of the immune response. *J. Leukoc. Biol.* **60**, 159–166 (1996). doi: 10.1002/jlb.60.2.159; pmid: 8773576
3. E. L. Yap, M. E. Greenberg, Activity-regulated transcription: Bridging the gap between neural activity and behavior. *Neuron* **100**, 330–348 (2018). doi: 10.1016/j.neuron.2018.10.013; pmid: 30359600
4. E. Benito, A. Barco, The neuronal activity-driven transcriptome. *Mol. Neurobiol.* **51**, 1071–1088 (2015). doi: 10.1007/s12035-014-8772-z; pmid: 24935719
5. S. Bahrami, F. Drabløs, Gene regulation in the immediate-early response process. *Adv. Biol. Regul.* **62**, 37–49 (2016). doi: 10.1016/j.jbior.2016.05.001; pmid: 27220739
6. G. Bossis, P. Ferrara, C. Acquaviva, I. Jariel-Encontre, M. Piechaczyk, c-Fos proto-oncoprotein is degraded by the proteasome independently of its own ubiquitinylation in vivo. *Mol. Cell. Biol.* **23**, 7425–7436 (2003). doi: 10.1128/MCB.23.20.7425-7436.2003; pmid: 14517309
7. T. Gomard *et al.*, Fos family protein degradation by the proteasome. *Biochem. Soc. Trans.* **36**, 858–863 (2008). doi: 10.1042/BST0360858; pmid: 18793151
8. H. C. Yen, Q. Xu, D. M. Chou, Z. Zhao, S. J. Elledge, Global protein stability profiling in mammalian cells. *Science* **322**, 918–923 (2008). doi: 10.1126/science.1160489; pmid: 18988847
9. M. E. Greenberg, E. B. Ziff, Stimulation of 3T3 cells induces transcription of the c-fos proto-oncogene. *Nature* **311**, 433–438 (1984). doi: 10.1038/311433a0; pmid: 6090941
10. J. Y. Joo, K. Schaukowitch, L. Farbiak, G. Kilaru, T. K. Kim, Stimulus-specific combinatorial functionality of neuronal c-fos enhancers. *Nat. Neurosci.* **19**, 75–83 (2016). doi: 10.1038/nn.4170; pmid: 26595656
11. A. R. Mardinly *et al.*, Sensory experience regulates cortical inhibition by inducing IGF1 in VIP neurons. *Nature* **531**, 371–375 (2016). doi: 10.1038/nature17187; pmid: 26958833
12. N. Sagehashi *et al.*, Insulin enhances gene expression of midnolin, a novel genetic risk factor for Parkinson's disease, via extracellular signal-regulated kinase, phosphoinositide 3-kinase and multiple transcription factors in SH-SY5Y cells. *J. Pharmacol. Exp. Ther.* **381**, 68–78 (2022). doi: 10.1124/jpet.121.001076; pmid: 35241633
13. I. Koren *et al.*, the eukaryotic proteome is shaped by E3 ubiquitin ligases targeting C-terminal degrons. *Cell* **173**, 1622–1635.e14 (2018). doi: 10.1016/j.cell.2018.04.028; pmid: 29779948
14. J. Du *et al.*, Stuxnet facilitates the degradation of polycomb protein during development. *Dev. Cell* **37**, 507–519 (2016). doi: 10.1016/j.devcel.2016.05.013; pmid: 27326929
15. H. W. Mittrücker *et al.*, Requirement for the transcription factor LSIRF/IRF4 for mature B and T lymphocyte function. *Science* **275**, 540–543 (1997). doi: 10.1126/science.275.5299.540; pmid: 8999800
16. U. Klein *et al.*, Transcription factor IRF4 controls plasma cell differentiation and class-switch recombination. *Nat. Immunol.* **7**, 773–782 (2006). doi: 10.1038/ni1357; pmid: 16767092
17. E. L. Huttlin *et al.*, Dual proteome-scale networks reveal cell-specific remodeling of the human interactome. *Cell* **184**, 3022–3040.e28 (2021). doi: 10.1016/j.cell.2021.04.011; pmid: 33961781
18. T. H. Wu, L. Shi, A. W. Lowe, M. R. Nicolls, P. N. Kao, Inducible expression of immediate early genes is regulated through dynamic chromatin association by NF45/ILF2 and NF90/NF110/ILF3. *PLOS ONE* **14**, e0216042 (2019). doi: 10.1371/journal.pone.0216042; pmid: 31022259
19. M. L. Hyer *et al.*, A small-molecule inhibitor of the ubiquitin activating enzyme for cancer treatment. *Nat. Med.* **24**, 186–193 (2018). doi: 10.1038/nm.4474; pmid: 29334375
20. A. Montagnoli *et al.*, Ubiquitination of p27 is regulated by Cdk-dependent phosphorylation and trimeric complex formation. *Genes Dev.* **13**, 1181–1189 (1999). doi: 10.1101/gad.13.9.1181; pmid: 10323868
21. L. M. Tsvetkov, K. H. Yeh, S. J. Lee, H. Sun, H. Zhang, p27(Kip1) ubiquitination and degradation is regulated by the SCF(Skp2) complex through phosphorylated Thr187 in p27. *Curr. Biol.* **9**, 661–664 (1999). doi: 10.1016/S0960-9822(99)80290-5; pmid: 10375532
22. M. Welcker *et al.*, Two diphosphorylated degrons control c-Myc degradation by the Fbw7 tumor suppressor. *Sci. Adv.* **8**, eabl7872 (2022). doi: 10.1126/sciadv.abl7872; pmid: 35089787
23. M. M. Peña, Y. Y. Xing, S. Koli, F. G. Berger, Role of N-terminal residues in the ubiquitin-independent degradation of human thymidylate synthase. *Biochem. J.* **394**, 355–363 (2006). doi: 10.1042/BJ20051479; pmid: 16259621
24. K. N. Swatek, D. Komander, Ubiquitin modifications. *Cell Res.* **26**, 399–422 (2016). doi: 10.1038/cr.2016.39; pmid: 27012465
25. V. Chau *et al.*, A multiubiquitin chain is confined to specific lysine in a targeted short-lived protein. *Science* **243**, 1576–1583 (1989). doi: 10.1126/science.2538923; pmid: 2538923
26. J. Jumper *et al.*, Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021). doi: 10.1038/s41586-021-03819-2; pmid: 34265844
27. N. Zheng, N. Shabek, Ubiquitin ligases: Structure, function, and regulation. *Annu. Rev. Biochem.* **86**, 129–157 (2017). doi: 10.1146/annurev-biochem-060815-014922; pmid: 28375744
28. I. Dikic, S. Wakatsuki, K. J. Walters, Ubiquitin-binding domains - from structures to functions. *Nat. Rev. Mol. Cell Biol.* **10**, 659–671 (2009). doi: 10.1038/nrm2767; pmid: 19773779
29. C. Schauber *et al.*, Rad23 links DNA repair to the ubiquitin/proteasome pathway. *Nature* **391**, 715–718 (1998). doi: 10.1038/35661; pmid: 9490418
30. S. Elsasser *et al.*, Proteasome subunit Rpn1 binds ubiquitin-like protein domains. *Nat. Cell Biol.* **4**, 725–730 (2002). doi: 10.1038/ncb845; pmid: 12198498
31. A. Bachmair, D. Finley, A. Varshavsky, In vivo half-life of a protein is a function of its amino-terminal residue. *Science* **234**, 179–186 (1986). doi: 10.1126/science.3018930; pmid: 3018930
32. A. Varshavsky, Naming a targeting signal. *Cell* **64**, 13–15 (1991). doi: 10.1016/0092-8674(91)90202-A; pmid: 1986863
33. R. T. Timms, I. Koren, Tying up loose ends: The N-degron and C-degron pathways of protein degradation. *Biochem. Soc. Trans.* **48**, 1557–1567 (2020). doi: 10.1042/BST20191094; pmid: 32627813
34. A. Varshavsky, N-degron and C-degron pathways of protein degradation. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 358–366 (2019). doi: 10.1073/pnas.1816596116; pmid: 30622213
35. E. J. Nestler, M. Barrot, D. W. Self, DeltaFosB: A sustained molecular switch for addiction. *Proc. Natl. Acad. Sci. U.S.A.* **98**, 11042–11046 (2001). doi: 10.1073/pnas.191352698; pmid: 11572966
36. B. T. Hope *et al.*, Induction of a long-lasting AP-1 complex composed of altered Fos-like proteins in brain by chronic cocaine and other chronic treatments. *Neuron* **13**, 1235–1244 (1994). doi: 10.1016/0896-6273(94)90061-2; pmid: 7946359
37. M. B. Kelz *et al.*, Expression of the transcription factor deltaFosB in the brain controls sensitivity to cocaine. *Nature* **401**, 272–276 (1999). doi: 10.1038/45790; pmid: 10499584
38. S. C. Li, N. K. Goto, K. A. Williams, C. M. Deber, Alpha-helical, but not beta-sheet, propensity of proline is determined by peptide environment. *Proc. Natl. Acad. Sci. U.S.A.* **93**, 6676–6681 (1996). doi: 10.1073/pnas.93.13.6676; pmid: 8692877
39. M. Tsukahara, H. Suemori, S. Noguchi, Z. S. Ji, H. Tsunoo, Novel nucleolar protein, midnolin, is expressed in the mesencephalon during mouse development. *Gene* **254**, 45–55 (2000). doi: 10.1016/S0378-1119(00)00259-6; pmid: 10974535
40. F. T. Gallo, C. Katche, J. F. Morici, J. H. Medina, N. V. Weisstaub, Immediate early genes, memory and psychiatric disorders: Focus on c-Fos, Egr1 and Arc. *Front. Behav. Neurosci.* **12**, 79 (2018). doi: 10.3389/fnbeh.2018.00079; pmid: 29755331
41. A. L. Shaffer *et al.*, IRF4 addiction in multiple myeloma. *Nature* **454**, 226–231 (2008). doi: 10.1038/nature07064; pmid: 18568025
42. P. S. Biswas, G. Bhagat, A. B. Pernis, IRF4 and its regulators: Evolving insights into the pathogenesis of inflammatory arthritis? *Immunol. Rev.* **233**, 79–96 (2010). doi: 10.1111/j.0105-2896.2009.00864.x; pmid: 20192994
43. R. Nayar *et al.*, TCR signaling via Tec kinase ITK and interferon regulatory factor 4 (IRF4) regulates CD8+ T-cell differentiation. *Proc. Natl. Acad. Sci. U.S.A.* **109**, E2794–E2802 (2012). doi: 10.1073/pnas.1205742109; pmid: 23011795
44. J. W. Williams *et al.*, Transcription factor IRF4 drives dendritic cells to promote Th2 differentiation. *Nat. Commun.* **4**, 2990 (2013). doi: 10.1038/ncomms3990; pmid: 24356538
45. J. Eguchi *et al.*, Transcriptional control of adipose lipid handling by IRF4. *Cell Metab.* **13**, 249–259 (2011). doi: 10.1016/j.cmet.2011.02.005; pmid: 21356515
46. Y. Murakami, S. Matsufuji, S. Hayashi, N. Tanahashi, K. Tanaka, Degradation of ornithine decarboxylase by the 26S proteasome. *Biochem. Biophys. Res. Commun.* **267**, 1–6 (2000). doi: 10.1006/bbrc.1999.1706; pmid: 10623564
47. L. Ghoda, T. van Daalen Wetters, M. Macrae, D. Ascherman, P. Coffino, Prevention of rapid intracellular degradation of ODC by a carboxyl-terminal truncation. *Science* **243**, 1493–1495 (1989). doi: 10.1126/science.2928784; pmid: 2928784
48. Y. Murakami *et al.*, Ornithine decarboxylase is degraded by the 26S proteasome without ubiquitination. *Nature* **360**, 597–599 (1992). doi: 10.1038/360597a0; pmid: 1334232
49. M. A. Hoyt, M. Zhang, P. Coffino, Ubiquitin-independent mechanisms of mouse ornithine decarboxylase degradation are conserved between mammalian and fungal cells. *J. Biol. Chem.* **278**, 12135–12143 (2003). doi: 10.1074/jbc.M211802200; pmid: 12562772
50. J. S. Merkel, J. M. Sturtevant, L. Regan, Sidechain interactions in parallel beta sheets: The energetics of cross-strand pairings. *Structure* **7**, 1333–1343 (1999). doi: 10.1016/S0969-2126(00)80023-4; pmid: 10574793
51. I. Marín, The ubiquilin gene family: Evolutionary patterns and functional insights. *BMC Evol. Biol.* **14**, 63 (2014). doi: 10.1186/1471-2148-14-63; pmid: 24674348

52. S. Raasi, C. M. Pickart, Rad23 ubiquitin-associated domains (UBA) inhibit 26 S proteasome-catalyzed proteolysis by sequestering lysine 48-linked polyubiquitin chains. *J. Biol. Chem.* **278**, 8951–8959 (2003). doi: 10.1074/jbc.M212841200; pmid: 12643283
53. L. Chen, U. Shinde, T. G. Ortolan, K. Madura, Ubiquitin-associated (UBA) domains in Rad23 bind ubiquitin and promote inhibition of multi-ubiquitin chain assembly. *EMBO Rep.* **2**, 933–938 (2001). doi: 10.1093/embo-reports/kve203; pmid: 11571271
54. K. K. Joshi, M. Bergé, S. K. Radhakrishnan, P. H. Viollier, P. Chien, An adaptor hierarchy regulates proteolysis during a bacterial cell cycle. *Cell* **163**, 419–431 (2015). doi: 10.1016/j.cell.2015.09.030; pmid: 26451486
55. N. J. Kuhlmann, D. Doxsey, P. Chien, Cargo competition for a dimerization interface restricts and stabilizes a bacterial protease adaptor. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2010523118 (2021). doi: 10.1073/pnas.2010523118; pmid: 33875581
56. N. Sharma *et al.*, ARNT2 tunes activity-dependent gene expression through NCoR2-mediated repression and NPAS4-mediated activation. *Neuron* **102**, 390–406.e9 (2019). doi: 10.1016/j.neuron.2019.02.007; pmid: 30846309
57. J. G. Doench *et al.*, Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat. Biotechnol.* **34**, 184–191 (2016). doi: 10.1038/nbt.3437; pmid: 26780180
58. R. T. Timms *et al.*, A glycine-specific N-degron pathway mediates the quality control of protein *N*-myristoylation. *Science* **365**, eaaw4912 (2019). doi: 10.1126/science.aaw4912; pmid: 31273098
59. T. D. Martin *et al.*, The adaptive immune system is a major driver of selection for tumor suppressor gene inactivation. *Science* **373**, 1327–1335 (2021). doi: 10.1126/science.abg5784; pmid: 34529489
60. M. Martin, Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. J.* **17**, 10–12 (2011). doi: 10.14806/ej.17.1.200
61. B. Langmead, S. L. Salzberg, Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012). doi: 10.1038/nmeth.1923; pmid: 22388286
62. W. Li *et al.*, MAGeCK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. *Genome Biol.* **15**, 554 (2014). doi: 10.1186/s13059-014-0554-4; pmid: 25476604
63. A. Subramanian *et al.*, Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 15545–15550 (2005). doi: 10.1073/pnas.0506580102; pmid: 16199517
64. A. Liberzon *et al.*, Molecular signatures database (MSigDB) 3.0. *Bioinformatics* **27**, 1739–1740 (2011). doi: 10.1093/bioinformatics/btr260; pmid: 21546393
65. Y. Lin *et al.*, Activity-dependent regulation of inhibitory synapse development by Npas4. *Nature* **455**, 1198–1204 (2008). doi: 10.1038/nature07319; pmid: 18815592
66. C. Nardone *et al.*, A central role for regulated protein stability in the control of TFE3 and MITF by nutrients. *Mol. Cell* **83**, 57–73.e9 (2023). doi: 10.1016/j.molcel.2022.12.013; pmid: 36608670
67. J. Martin *et al.*, Protein secondary structure assignment revisited: A detailed analysis of different assignment methods. *BMC Struct. Biol.* **5**, 17 (2005). doi: 10.1186/1472-6807-5-17; pmid: 16164759
68. W. Kabsch, C. Sander, Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **22**, 2577–2637 (1983). doi: 10.1002/bip.360221211; pmid: 6667333
69. K. Tunyasuvunakool *et al.*, Highly accurate protein structure prediction for the human proteome. *Nature* **596**, 590–596 (2021). doi: 10.1038/s41586-021-03828-1; pmid: 34293799
70. O. D. Monera, T. J. Sereda, N. E. Zhou, C. M. Kay, R. S. Hodges, Relationship of sidechain hydrophobicity and alpha-helical propensity on the stability of the single-stranded amphipathic alpha-helix. *J. Pept. Sci.* **1**, 319–329 (1995). doi: 10.1002/psc.310010507; pmid: 9223011
71. T. J. Sereda, C. T. Mant, F. D. Sönnichsen, R. S. Hodges, Reversed-phase chromatography of synthetic amphipathic alpha-helical peptides as a model for ligand/receptor interactions. Effect of changing hydrophobic environment on the relative hydrophilicity/hydrophobicity of amino acid side-chains. *J. Chromatogr. A* **676**, 139–153 (1994). doi: 10.1016/0021-9673(94)00371-8; pmid: 7921171
72. G. A. Collins, A. L. Goldberg, Proteins containing ubiquitin-like (Ubl) domains not only bind to 26S proteasomes but also induce their activation. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 4664–4674 (2020). doi: 10.1073/pnas.1915534117; pmid: 32071216
73. X. Gu *et al.*, Data for: The midnolin-proteasome pathway catches proteins for ubiquitination-independent degradation, Dryad (2023); https://doi.org/10.5061/dryad.m905qfv6g.

ACKNOWLEDGMENTS

We thank all members of the Greenberg and Elledge labs for helpful suggestions and contributions, especially T. Whitwam, O. Dagliyan, and L. Hu for the generous gift of the NR4A1 antibody and I. Vogel, E. Assad, and N. Pajarillo for help with primary cortical neuron dissection. We also thank T. Martin for help with the midnolin knock-in strategy and M. Dezfulian for sharing unpublished expression vectors; J. Nelson and the Harvard Medical School Department of Immunology Flow Cytometry Core Facility for invaluable help in cell sorting the genetic screens; the Research Computing Group at Harvard Medical School as AlphaFold analyses were performed on the O2 High Performance Compute Cluster; B. Palanski and P. Cole for access and help with mass spectrometry; and A. Goldberg, S. Shao, and M. Negasi for valuable suggestions for our manuscript. Raw data for this manuscript can be accessed using Dryad (*73*). **Funding:** This work was supported by the National Institutes of Health (NIH grant R01 NS115965 to M.E.G.; and National Institute on Aging grant AG11085 to S.J.E.), X.G. is the National Mah Jongg League Fellow of the Damon Runyon Cancer Research Foundation (DRG-2469-22). C.N. is supported by the National Science Foundation Graduate Research Fellowship program. N.K. is supported by funding from the NIH (grant T32 HG002295). S.J.E. is an investigator with the Howard Hughes Medical Institute. The Greenberg Laboratory is supported by the Allen Discovery Center Program, a Paul G. Allen Frontiers Group advised program of the Paul G. Allen Family Foundation, and the Tang-Yang Autism Center at Harvard Medical School. **Author contributions:** The study was conceptualized by X.G. and C.N. Wet and dry experiments were performed by X.G., C.N., N.K., and A.M. Funding was secured by S.J.E. and M.E.G. The study was supervised by S.J.E. and M.E.G. The manuscript was written by X.G., C.N., S.J.E., and M.E.G. **Competing interests:** S.J.E. is a founder of TSCAN Therapeutics, MAZE Therapeutics, ImmuneID, and Mirimus serves on the scientific advisory boards of Homology Medicines, ImmuneID, MAZE Therapeutics, X-Chem, and TSCAN Therapeutics; and is an adviser for MPM Capital. The remaining authors declare no competing interests. **Data and materials availability:** Any reagents that are unique to this study will be made available upon request. Requests for reagents or any additional information necessary to reanalyze the data reported in this paper should be directed to and will be fulfilled by the corresponding authors. The raw DNA sequencing data, mass spectrometry data, and PDB files for the midnolin-substrate AlphaFold predictions were deposited to Dryad (*73*).

License information: Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse This article is subject to HHMI's Open Access to Publications policy. HHMI lab heads have previously granted a nonexclusive CC BY 4.0 license to the public and a sublicensable license to HHMI in their research articles. Pursuant to those licenses, the author-accepted manuscript (AAM) of this article can be made freely available under a CC BY 4.0 license immediately upon publication.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adh5021
Figs. S1 to S11
References
Data S1 to S4
MDAR Reproducibility Checklist

Submitted 8 March 2023; resubmitted 1 June 2023
Accepted 11 July 2023
10.1126/science.adh5021

RESEARCH ARTICLE

METABOLISM

The gut microbiota reprograms intestinal lipid metabolism through long noncoding RNA *Snhg9*

Yuhao Wang[1,2,3,4]*, Meng Wang[1], Jiaxin Chen[5], Yun Li[4], Zheng Kuang[4]†, Chaitanya Dende[4], Prithvi Raj[4], Gabriella Quinn[4], Zehan Hu[4]‡, Tarun Srinivasan[4], Brian Hassell[4], Kelly A. Ruhn[4], Cassie L. Behrendt[4], Tingbo Liang[1,3], Xiaobing Dou[2], Zhangfa Song[6], Lora V. Hooper[4,7]*

The intestinal microbiota regulates mammalian lipid absorption, metabolism, and storage. We report that the microbiota reprograms intestinal lipid metabolism in mice by repressing the expression of long noncoding RNA (lncRNA) *Snhg9* (small nucleolar RNA host gene 9) in small intestinal epithelial cells. *Snhg9* suppressed the activity of peroxisome proliferator–activated receptor γ (PPARγ)—a central regulator of lipid metabolism—by dissociating the PPARγ inhibitor sirtuin 1 from cell cycle and apoptosis protein 2 (CCAR2). Forced expression of *Snhg9* in the intestinal epithelium of conventional mice impaired lipid absorption, reduced body fat, and protected against diet-induced obesity. The microbiota repressed *Snhg9* expression through an immune relay encompassing myeloid cells and group 3 innate lymphoid cells. Our findings thus identify an unanticipated role for a lncRNA in microbial control of host metabolism.

The intestinal microbiota has an important impact on mammalian metabolism. In mice, resident intestinal bacteria enhance digestion of dietary polysaccharides (*1*), promote dietary lipid absorption by intestinal epithelial cells (*2*–*5*), and alter the function of adipose tissue (*6*–*9*). Consistent with the findings in mice, alterations in the composition of the human intestinal microbiota are associated with metabolic disorders, including obesity, type 2 diabetes, and cardiovascular disease (*10*–*12*). Given the rapidly increasing worldwide prevalence of metabolic disease (*13*), there is a pressing need to understand the mechanisms by which the intestinal microbiota affects host metabolism.

[1]Zhejiang Provincial Key Laboratory of Pancreatic Disease of The First Affiliated Hospital, Institute of Translational Medicine, Zhejiang University School of Medicine, Hangzhou, Zhejiang 310029, China. [2]School of Life Science, Zhejiang Chinese Medical University, Hangzhou, Zhejiang 310053, China. [3]Cancer Center, Zhejiang University, Hangzhou, Zhejiang 310029, China. [4]Department of Immunology, The University of Texas Southwestern Medical Center, Dallas, TX 75390, USA. [5]Department of Breast Surgery and Key Laboratory of Tumor Microenvironment and Immune Therapy of Zhejiang Province, Second Affiliated Hospital, Zhejiang University, Hangzhou, Zhejiang 310009, China. [6]Department of Colorectal Surgery and Key Laboratory of Biotherapy of Zhejiang Province, Sir Run Shaw Hospital, Zhejiang University, Hangzhou, Zhejiang 310016, China. [7]The Howard Hughes Medical Institute, The University of Texas Southwestern Medical Center, Dallas, TX 75390, USA.
***Corresponding author. Email: lora.hooper@utsouthwestern.edu (L.V.H.); yuhaowang@zju.edu.cn (Y.W.)**
†Present address: Department of Biological Sciences, Carnegie Mellon University, Pittsburgh, PA 15213, USA.
‡Present address: State Key Laboratory of Microbial Metabolism, Joint International Research Laboratory of Metabolic and Developmental Sciences, School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai 200240, China.

Long noncoding RNAs (lncRNAs) are RNA transcripts that are not translated but are nevertheless biologically functional. lncRNAs regulate biological processes, such as cell proliferation, cell death, tumorigenesis, and immunity (*14*–*17*), but little is known about their involvement in the regulation of host metabolism by the gut microbiota. We therefore investigated whether the presence of the gut microbiota affects the transcription of lncRNAs in small intestinal epithelial cells, which are central to microbial modulation of lipid absorption and metabolism (*2*–*5*).

Expression of lncRNA *Snhg9* is repressed by the microbiota

Whole-transcriptome sequencing [RNA sequencing (RNA-seq)] of small intestinal epithelial cells from conventionally raised (conventional) and germ-free mice identified 60 differentially expressed non–protein-coding genes, including 42 genes that encode lncRNAs (Fig. 1A and fig. S1, A and B). In particular, the lncRNA encoded by *Snhg9* (small nucleolar RNA host gene 9) showed reduced abundance in epithelial cells from conventional as compared with germ-free mice (Fig. 1B). Although lncRNAs are generally defined as >200 nucleotides (*18*) and mouse *Snhg9* RNA is only 183 nucleotides long, human *SNHG9* RNA is 233 nucleotides long, and thus both the mouse and human RNAs are designated as lncRNAs (*19*).

Snhg9 transcripts were detected in small intestinal epithelial cells of antibiotic-treated mice by in situ hybridization (fig. S2A). Analysis of published single-cell RNA-seq data from mouse intestinal epithelial cells indicated that *Snhg9* transcripts were mostly from stem cells, enterocytes, and enteroendocrine cells (a specialized epithelial cell lineage) (*20*) (fig. S
Using quantitative polymerase chain reaction (qPCR), we confirmed that epithelial cells from conventional mice express less *Snhg9* RNA when compared with epithelial cells from both germ-free and antibiotic-treated mice (Fig. 1C). Thus, *Snhg9* expression in small intestinal epithelial cells is reduced in the presence of the microbiota.

lncRNA *Snhg9* binds to CCAR2

We next sought to illuminate the biological functions of *Snhg9*. Because other lncRNAs bind to and regulate the activity of proteins (*21*–*23*), we screened for *Snhg9*-protein interactions in small intestinal epithelial cells. We performed an RNA-protein pull-down assay in epithelial cell lysates using in vitro transcribed *Snhg9* or antisense *Snhg9*. Mass spectrometry identified the most abundant interacting protein as cell cycle and apoptosis protein 2 (CCAR2; also known as deleted in breast cancer 1, or DBC1) (*24*) (Fig. 2, A and B). This finding was supported by immunoblot detection of CCAR2 among the proteins precipitated by *Snhg9* (Fig. 2C).

To test whether *Snhg9* binds directly to CCAR2, we performed pull-down assays with in vitro transcribed *Snhg9* and recombinant CCAR2. Consistent with our findings in epithelial cell lysates, recombinant CCAR2 was precipitated by *Snhg9* RNA but not by antisense *Snhg9* RNA or polyA RNA (Fig. 2D), which indicates that CCAR2 binds directly to *Snhg9* RNA. Additionally, mutations in *Snhg9* affected binding to CCAR2. Deletion of 28 nucleotides in the middle of the *Snhg9* sequence did not affect *Snhg9* RNA binding to CCAR2. However, binding was reduced by 24-nucleotide deletions in the 3′ and 5′ regions, which target a predicted loop in the *Snhg9* secondary structure (Fig. 2E and fig. S3, A and B). These data support a direct binding interaction and indicate that *Snhg9* is a protein-binding lncRNA that binds directly to CCAR2.

lncRNA *Snhg9* dissociates CCAR2 from the PPARγ inhibitor SIRT1, repressing PPARγ activity

We next investigated the cellular and physiological consequences of lncRNA *Snhg9* binding to CCAR2. CCAR2 is an endogenous inhibitor of the deacetylase sirtuin 1 (SIRT1) (*25*, *26*). Among its several functions, SIRT1 regulates lipid metabolism by interacting with the transcription factor peroxisome proliferator–activated receptor gamma (PPARγ), which directs the transcription of lipid metabolic genes (*27*–*29*). SIRT1 represses PPARγ activity through two mechanisms—by deacetylating PPARγ or by docking with nuclear receptor corepressor 1 (NcoR1), a PPARγ cofactor. Both mechanisms decrease *Pparg* expression and reduce lipid metabolism (*30*, *31*). We reasoned that by binding to CCAR2, *Snhg9*

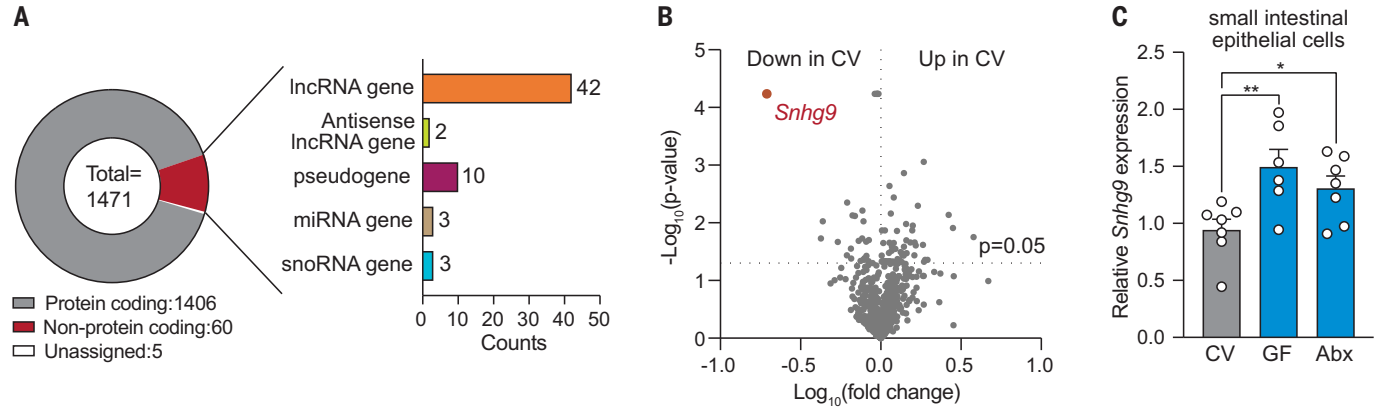


Fig. 1. Expression of lncRNA *Snhg9* is repressed by the microbiota. (**A**) Whole-transcriptome sequencing of small intestinal epithelial cells recovered by laser capture microdissection from conventional and germ-free mice. Genes differentially expressed between conventional and germ-free mice. *Snhg9* is highlighted in red. (**C**) qPCR analysis of *Snhg9* expression in small intestinal epithelial cells recovered by laser capture microdissection from conventional (CV), germ-free (GF), and antibiotic-treated (Abx) mice. Results are representative of at least two independent experiments. Means ± SEMs are plotted; each data point represents one mouse. **P* < 0.05; ***P* < 0.01; two-tailed Student's *t* test.

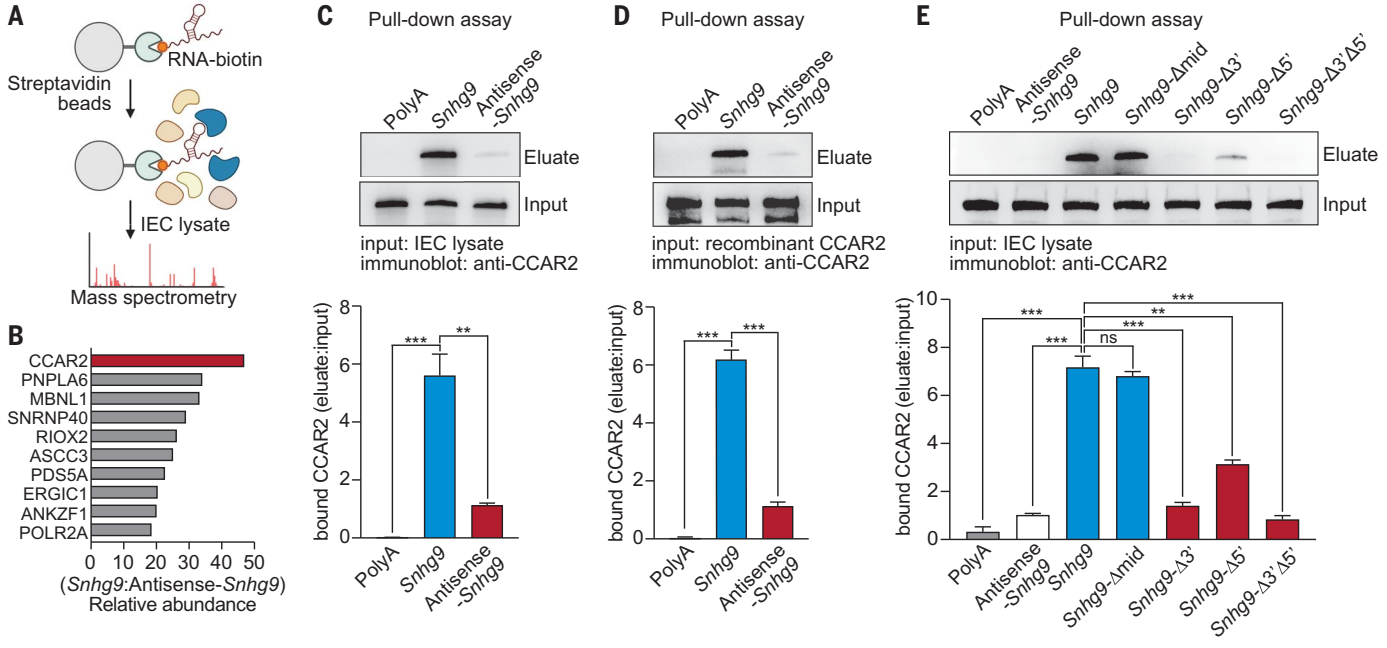


Fig. 2. lncRNA *Snhg9* binds to CCAR2. (**A**) Schematic of RNA-protein pull-down assays in small intestinal epithelial cell (IEC) lysates in combination with mass spectrometry analysis. (Created using BioRender.) (**B**) Ten most abundant *Snhg9*-binding proteins identified by mass spectrometry. CCAR2 is highlighted in red. (**C**) (Top) Representative immunoblot of CCAR2 in proteins pulled down from small intestinal IEC lysates by polyA RNA (negative control), *Snhg9*, or antisense *Snhg9*. (Bottom) Band intensities were quantified by densitometry and normalized to input. *N* = 3 experimental replicates per group. (**D**) (Top) Representative immunoblot of recombinant CCAR2 pulled down by polyA RNA, *Snhg9*, or antisense *Snhg9*. (Bottom) Intensities were quantified by densitometry and normalized to input. *N* = 3 experimental replicates per group. (**E**) (Top) Representative immunoblot of CCAR2 in proteins pulled down from small intestinal IEC lysates by polyA RNA, antisense *Snhg9*, *Snhg9*, *Snhg9* with 28 nucleotides deleted from the middle of the sequence (*Snhg9*-Δmid), *Snhg9* with 3′-deletion of 24 nucleotides (*Snhg9*-Δ3′), *Snhg9* with 5′-deletion of 24 nucleotides (*Snhg9*-Δ5′), or *Snhg9* with both 3′- and 5′-deletion of 24 nucleotides (*Snhg9*-Δ3′Δ5′). (Bottom) Band intensities were quantified by densitometry and normalized to input. *N* = 3 experimental replicates per group. All experiments are representative of at least two independent experiments. Means ± SEMs are plotted. ***P* < 0.01; ****P* < 0.001; ns, not significant; two-tailed Student's *t* test.

might inhibit its interaction with SIRT1, thereby rescuing SIRT1 deacetylase and NcoR1 binding activities from CCAR2 inhibition. In fact, overexpression of *Snhg9* in HEK-293T cells largely abolished CCAR2 binding to SIRT1 (Fig. 3, A and B) and increased both SIRT1 deacetylase activity (Fig. 3C) and binding to NcoR1 (Fig. 3, D and E). These data indicate that *Snhg9* promotes SIRT1 activity by sequestering the inhibitory protein CCAR2.

Because *Snhg9* rescued SIRT1 activity from CCAR2 inhibition, we predicted that *Snhg9*

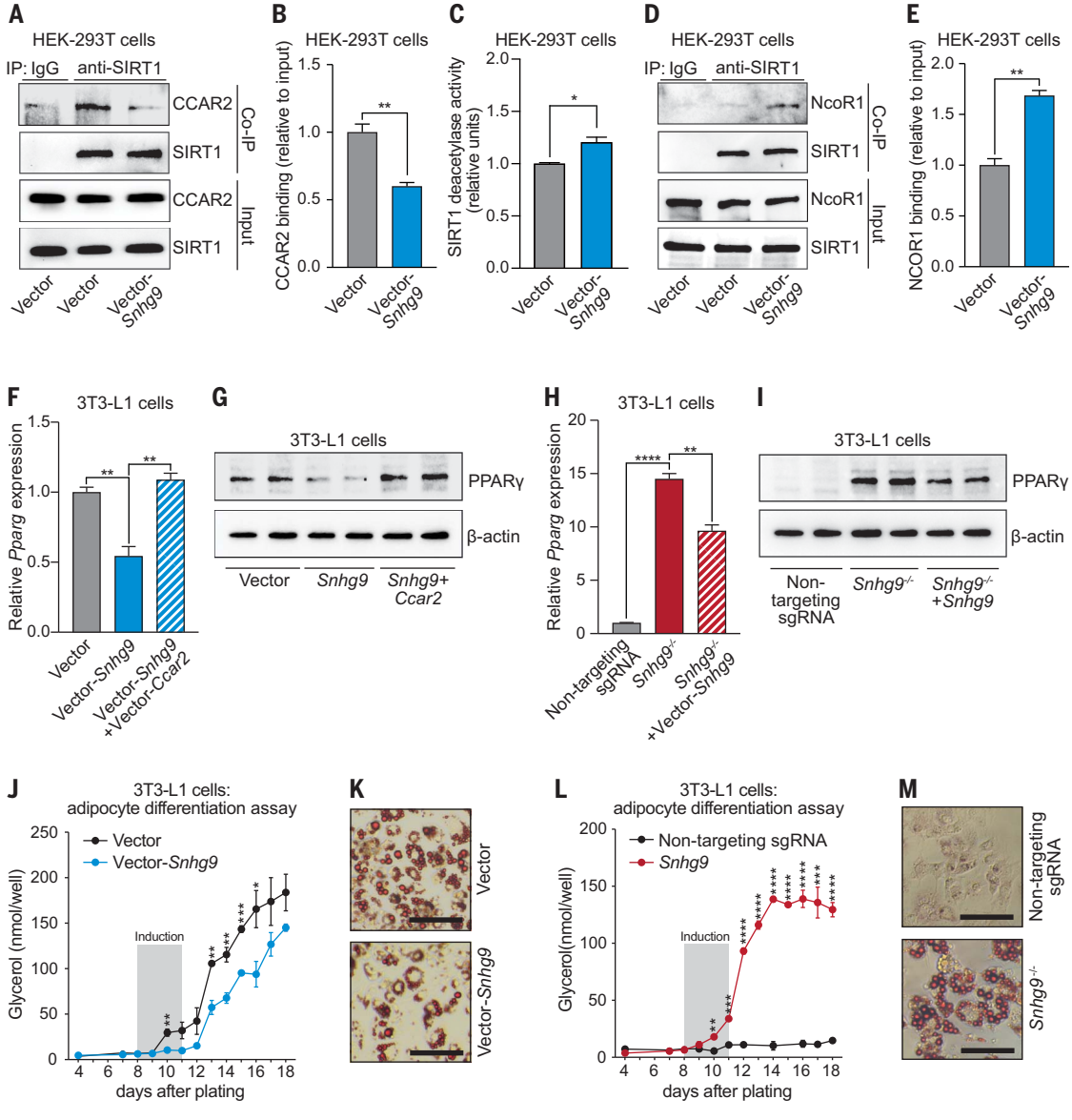


Fig. 3. lncRNA *Snhg9* dissociates CCAR2 from the PPARγ inhibitor SIRT1, repressing PPARγ activity. (**A**) Coimmunoprecipitation (co-IP) of CCAR2 and SIRT1 with anti-SIRT1 antibody or immunoglobulin G (IgG) isotype control. HEK-293T cells were transfected with empty vector or *Snhg9*-encoding vector. Proteins were detected by immunoblot. (**B**) Band intensities in (A) were quantified by densitometry and normalized to input. $N = 3$ experimental replicates per group. (**C**) Relative SIRT1 deacetylase activity in HEK-293T cells transfected with empty vector or *Snhg9*-encoding vector. $N = 5$ experimental replicates per group. (**D**) Coimmunoprecipitation of NcoR1 and SIRT1 with anti-SIRT1 antibody or IgG isotype control. HEK-293T cells were transfected with empty vector or *Snhg9*-encoding vector. Proteins were detected by immunoblot. (**E**) Band intensities in (D) were quantified by densitometry and normalized to input. $N = 3$ experimental replicates per group. (**F**) qPCR analysis of *Pparg* expression in 3T3-L1 cells with stable expression of *Snhg9* or coexpression of *Snhg9* and *Ccar2*. Cells were transduced with empty vector as a control. $N = 4$ experimental replicates per group. (**G**) Immunoblot detection of PPARγ and β-actin (control) in 3T3-L1 cells from (F). (**H**) qPCR analysis of *Pparg* expression in *Snhg9*$^{-/-}$ 3T3-L1 cells that were untreated or rescued by *Snhg9* expression and in cells edited with nontargeting sgRNA. $N = 4$ experimental replicates per group. (**I**) Immunoblot detection of PPARγ and β-actin (control) in 3T3-L1 cells from (H). (**J**) *Snhg9* was stably expressed in 3T3-L1 cells, and their differentiation to adipocytes was assessed by measuring glycerol as a readout of triglyceride accumulation. Cells were transduced with empty vector as a control. $N = 5$ experimental replicates per group. (**K**) Lipids were detected by Oil Red O staining of differentiated cells from (J). Scale bars, 30 μm. (**L**) *Snhg9*$^{-/-}$ 3T3-L1 cells and cells edited with nontargeting sgRNA were assessed for differentiation to adipocytes as in (J). $N = 5$ experimental replicates per group. Note that the multiple cell passages required by the CRISPR mutant selection process result in suppression of *Pparg* expression in the cells edited with nontargeting sgRNA (*32*). (**M**) Lipids were detected by Oil Red O staining of differentiated cells from (L). Scale bars, 30 μm. All experiments are representative of at least two independent experiments. Means ± SEMs are plotted. $^{*}P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$; $^{****}P < 0.0001$; two-tailed Student's t test.

would repress PPARγ expression and activity. We therefore evaluated the impact of *Snhg9* on *Pparg* expression and lipid metabolism in cells. PPARγ controls differentiation of the mouse fibroblast cell line 3T3-L1 into adipocytes after chemical induction (*30*). We generated 3T3-L1 cells that stably express *Snhg9* (fig. S4A) and confirmed binding of *Snhg9* to CCAR2 by RNA immunoprecipitation (RIP) assay (fig. S4B). Stable expression of *Snhg9* inhibited expression of *Pparg* and its protein product PPARγ, whereas coexpression of *Ccar2* rescued expression of both *Pparg* and PPARγ (Fig. 3, F and G). Accordingly, there was reduced transcription of PPARγ-controlled genes, including *Cd36* (encoding a fatty acid transporter), *Fabp4* (encoding a fatty acid binding protein), and *Lpl* (encoding lipoprotein lipase) (fig. S4, C to E). Similarly, stable expression of *Snhg9* in mouse small intestinal organoids reduced expression of *Pparg* and PPARγ-controlled genes (fig. S4, F to J). These effects were not the result of expression of the putative *Snhg9* open reading frame (ORF) (fig. S5A) because expression of the ORF alone failed to reduce expression of *Pparg* and its target genes, and *Snhg9* with mutated start and stop codons retained the ability to reduce *Pparg* expression (fig. S5B). Together, these results indicate that *Snhg9* represses PPARγ expression by binding to CCAR2.

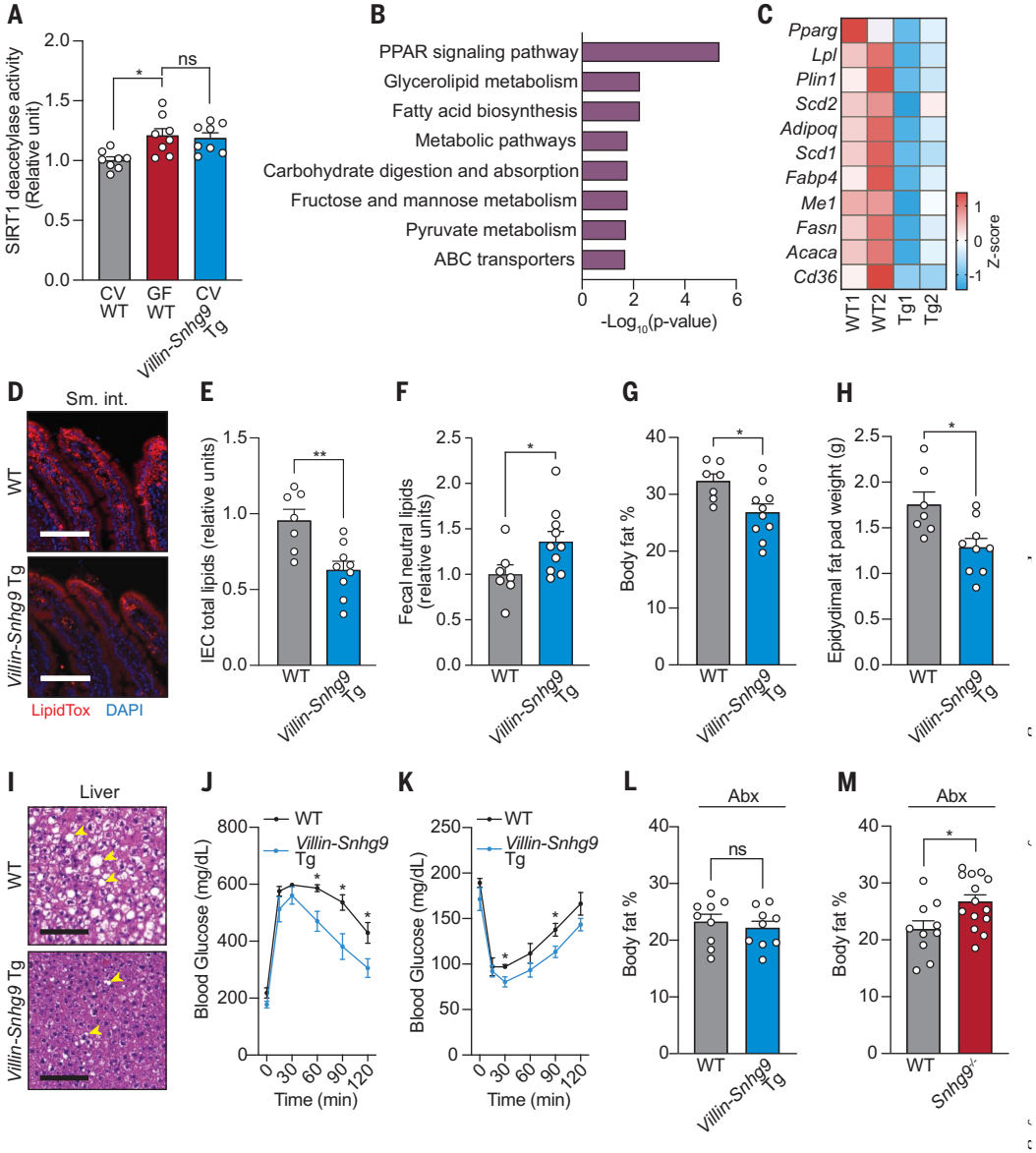


Fig. 4. *Villin-Snhg9* transgenic mice have reduced lipid absorption and are protected from high-fat diet–induced metabolic disorders. (**A**) Relative SIRT1 deacetylation activity in IECs from conventional wild-type, germ-free wild-type, and conventional *Villin-Snhg9* transgenic (Tg) mice. (**B**) RNA-seq of intestines of wild-type and *Villin-Snhg9* Tg littermates. KEGG pathway analysis identifies pathways affected by *Snhg9* overexpression. ABC, ATP-binding cassette. (**C**) Heatmap visualizing expression levels of selected lipid metabolic genes with altered expression in the small intestines of wild-type (WT) and *Villin-Snhg9* Tg littermates. (**D**) LipidTox detection of fatty acids in the small intestines (sm. int.) of wild-type and *Villin-Snhg9* Tg littermates fed a high-fat diet. Scale bars, 100 μm. (**E**) Relative total lipid concentrations in isolated IECs from wild-type and *Villin-Snhg9* Tg littermates fed a high-fat diet. (**F**) Relative total neutral lipid concentrations in the feces of wild-type and *Villin-Snhg9* Tg littermates fed a high-fat diet. (**G** to **I**) Wild-type and *Villin-Snhg9* Tg littermates were fed a high-fat diet for 10 weeks and were assessed for body fat percentage (G), epididymal fat pad weight (H), and liver fat accumulation (examples are indicated with arrowheads) as indicated by hematoxylin and eosin staining (scale bars, 100 μm) (I). (**J** and **K**) Wild-type and *Villin-Snhg9* Tg littermates fed a high-fat diet were assessed for glucose tolerance (J) and insulin tolerance (K). $N = 5$ mice per group. (**L**) Body fat percentages of wild-type and *Villin-Snhg9* Tg littermates that were treated with antibiotics after switching to a high-fat diet for 10 weeks. (**M**) Body fat percentages of wild-type and *Snhg9*$^{-/-}$ littermates that were treated with antibiotics after switching to a high-fat diet for 10 weeks. All experiments are representative of at least two independent experiments. Means ± SEMs are plotted; each data point represents one mouse. *$P < 0.05$; **$P < 0.01$; ns, not significant; two-tailed Student's t test.

Further supporting this idea, inactivation of *Snhg9* in 3T3-L1 cells (*Snhg9*$^{-/-}$ cells) by CRISPR-Cas9 genome editing (fig. S6, A to D) increased expression of *Pparg*, PPARγ, and downstream target genes (Fig. 3, H and I, and fig. S6E). This increase occurred despite the high frequency of cell passage during the selection process, which tends to suppress *Pparg* expression. By contrast, cells edited with a nontargeting single-guide RNA (sgRNA) showed minimal *Pparg* expression as a result of a high number of cell passages (*32*) (Fig. 3, H and I), and reexpression of *Snhg9* in *Snhg9*$^{-/-}$ 3T3-L1 cells largely reversed *Pparg* expression (Fig. 3, H and I). Notably, overexpression of *Snora78*, which is flanked by the *Snhg9* exons, did not alter *Pparg* expression (fig. S6F), which suggests that *Snora78* deletion did not cause the increased expression of *Pparg* in *Snhg9*$^{-/-}$ cells, and deleting the *Snhg9* locus did not interfere with the expression of a nearby gene *Rps2* (fig. S6G). These data support the idea that *Snhg9* releases SIRT1 from CCAR2 inhibition.

Consistent with the reduced *Pparg* expression, *Snhg9* overexpression restrained the differentiation of 3T3-L1 cells to adipocytes (Fig. 3, J and K). Conversely, *Snhg9*$^{-/-}$ 3T3-L1 cells maintained the ability to differentiate and form lipid droplets (Fig. 3, L and M). However, cells edited with nontargeting sgRNA lost the ability to differentiate as a result of insufficient expression of *Pparg* arising from multiple cell passages (*32*) (Fig. 3, L and M). Taken together, our results reveal that lncRNA *Snhg9* inhibits *Pparg* expression and lipid metabolism by dissociating the CCAR2-SIRT1 complex.

Villin-Snhg9 transgenic mice have reduced lipid absorption

The presence of the intestinal microbiota enhances dietary lipid absorption and promotes obesity in mice fed a high-fat diet. Germ-free mice, which are microbiologically sterile and thus lack an intestinal microbiota, tend to absorb less lipid compared with conventional mice and thus are largely protected from high-fat diet–induced obesity (*2*, *3*, *7*, *8*). Because *Snhg9* suppresses lipid metabolism in cells (Fig. 3)

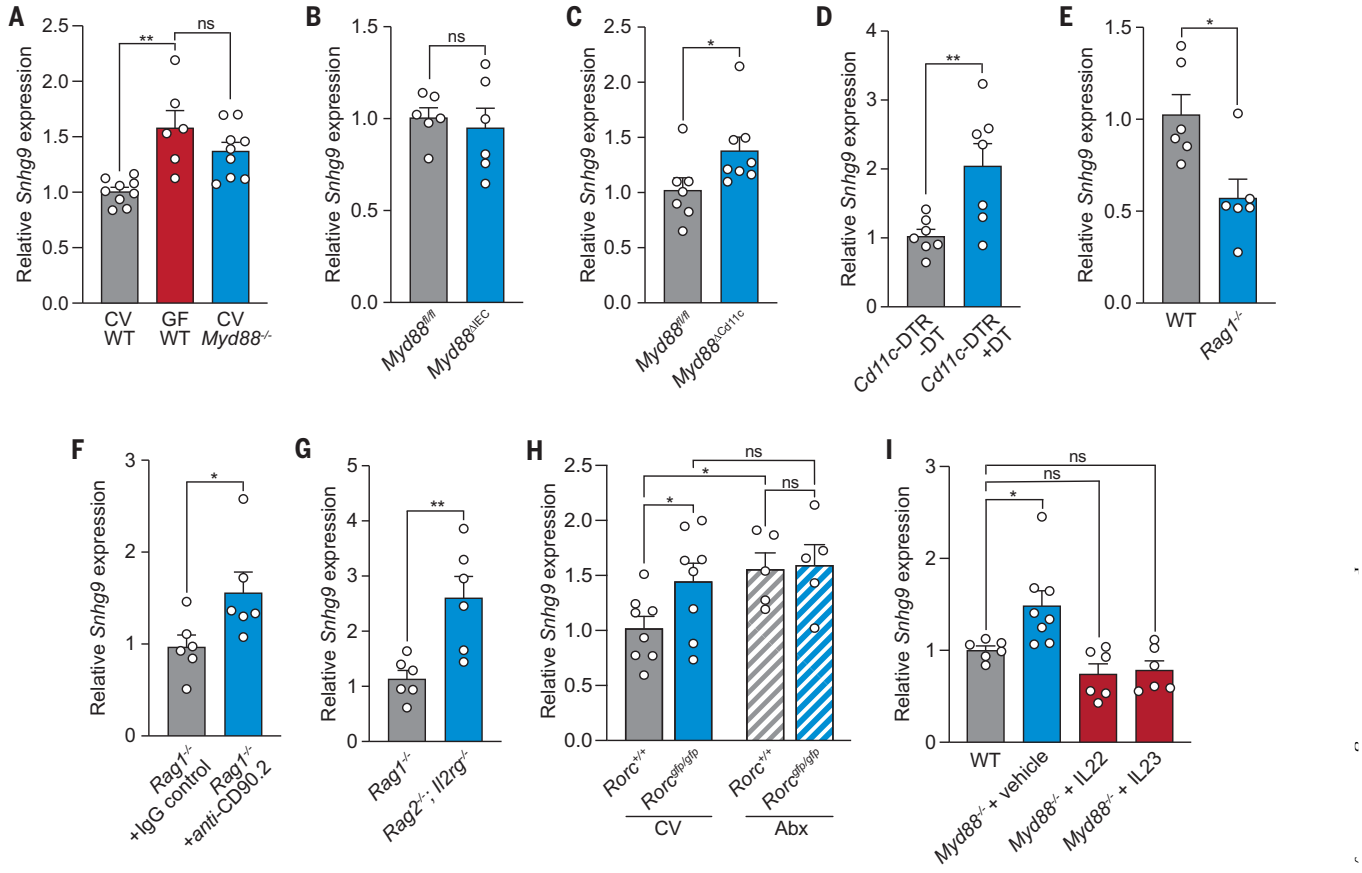


Fig. 5. The microbiota suppresses *Snhg9* expression through a myeloid cell–ILC3 relay. (**A** to **I**) *Snhg9* expression was measured by qPCR analysis of the small intestines of conventional (CV) wild-type, germ-free (GF) wild-type, and conventional *Myd88*$^{-/-}$ mice (A); *Myd88*$^{fl/fl}$ and *Myd88*$^{\Delta IEC}$ (epithelial cell–specific knockout) mice (B); *Myd88*$^{fl/fl}$ and *Myd88*$^{\Delta Cd11c}$ (*Cd11c*$^{+}$ cell–specific knockout) mice (C); *Cd11c*-DTR mice untreated or treated with *Diphtheria* toxin (DT) (D); wild-type (WT) and *Rag1*$^{-/-}$ mice (E); *Rag1*$^{-/-}$ mice injected using the intraperitoneal route with anti-CD90.2 antibody or IgG isotype control (F); *Rag1*$^{-/-}$ and *Rag2*$^{-/-}$;*Il2rg*$^{-/-}$ mice (G); *Rorc*$^{+/+}$ and *Rorc*$^{gfp/gfp}$ mice that were untreated (CV) or treated with antibiotics (Abx) (H); and *Myd88*$^{-/-}$ mice treated with recombinant IL-22, IL-23, or vehicle (I). All experiments are representative of at least two independent experiments. Means ± SEMs are plotted; each data point represents one mouse. *P < 0.05; **P < 0.01; ns, not significant; two-tailed Student's t test.

and is up-regulated in the intestines of germ-free mice (Fig. 1C), we hypothesized that *Snhg9* limits body fat accumulation in mice.

To test this hypothesis, we used the *Villin* promoter to force the expression of *Snhg9* in intestinal epithelial cells of conventional mice (*Villin-Snhg9* transgenic mice) (fig. S7, A to D). Consistent with our findings in cultured cells, SIRT1 deacetylase activity was higher in intestinal epithelial cells from conventional *Villin-Snhg9* transgenic mice compared with those from wild-type littermates, with activity levels comparable to those of wild-type germ-free mice (Fig. 4A).

To investigate whether *Snhg9* overexpression affects lipid metabolism in vivo, we compared the small intestinal transcriptomes of *Villin-Snhg9* transgenic mice with those of their wild-type littermates using RNA-seq. KEGG pathway analysis confirmed that *Snhg9* overexpression suppressed the expression of genes involved in metabolic pathways, including the PPAR signaling pathway (Fig. 4B). *Villin-Snhg9* transgenic mice showed reduced expression not only of *Pparg* but also genes involved in fatty acid absorption (such as *Cd36*), transport (such as *Fabp4*), and synthesis (such as *Scd1*) (Fig. 4C). Expression of these genes and their protein products was also reduced in germ-free mice compared with conventional wild-type mice (fig. S7, E and F). Consequently, *Villin-Snhg9* transgenic mice were similar to germ-free mice in that they had less lipid in their intestinal epithelial cells and more in their feces when compared with wild-type littermates (Fig. 4, D to F).

Villin-Snhg9 transgenic mice are protected from high-fat diet–induced metabolic disorders

When fed a normal chow diet, *Villin-Snhg9* transgenic mice had body weights similar to those of their wild-type littermates (fig. S7G). However, their body fat percentages and epididymal fat pad weights were reduced (fig. S7, H and I), and they were more glucose tolerant (fig. S7J). When switched to a high-fat diet for 10 weeks, *Villin-Snhg9* transgenic mice exhibited decreased body weights (fig. S7K), lower overall body fat percentages (Fig. 4G), smaller epididymal fat pads (Fig. 4H and fig. S7L), and milder liver steatosis (Fig. 4I) compared with their wild-type littermates. They also had lower serum triglycerides and free fatty acids (fig. S7, M and N), increased glucose tolerance, and decreased insulin resistance (Fig. 4, J and K). These phenotypes did not result from altered food intake, physical activity, respiratory exchange ratio, or microbiota composition (fig. S8, A to C).

Because the expression of intestinal *Snhg9* is suppressed by the microbiota (Fig. 1C), we further assessed the requirement for the microbiota in *Snhg9*-regulated lipid metabolism. When we depleted the microbiota by antibiotic treatment, wild-type mice fed a high-fat diet had lowered body fat percentages, similar to those of *Villin-Snhg9* transgenic littermates (Fig. 4L). This is consistent with the increased *Snhg9* expression in antibiotic-treated mice

(Fig. 1C). For comparison, we generated *Snhg9*$^{-/-}$ mice by CRISPR-Cas9–mediated gene targeting (fig. S9, A and B). Although the *Snhg9*$^{-/-}$ mice had body weights and body fat percentages similar to those of wild-type littermates when fed a normal chow diet (fig. S9, C and D), they had increased body fat percentages and weight gain when fed a high-fat diet, even when their microbiota were depleted with antibiotics (Fig. 4M and fig. S9E). These data support the conclusion that the microbiota promotes body fat accumulation in part by repressing intestinal *Snhg9* expression.

The microbiota suppresses *Snhg9* expression through a myeloid cell–ILC3 relay

Bacteria activate intestinal epithelial cell gene expression through Toll-like receptors (TLRs) and their common signaling adaptor MyD88 (*33*). Expression of intestinal *Snhg9* was increased in *Myd88*$^{-/-}$ mice compared with wild-type controls (Fig. 5A), which suggests that MyD88 is required for microbial repression of intestinal *Snhg9* expression. Although epithelial cell *Myd88* was dispensable for repression of *Snhg9* expression (Fig. 5B), mice with *Myd88* selectively deleted in CD11c^{+} cells showed elevated *Snhg9* expression (Fig. 5C), which suggests a role for CD11c^{+} cells in repressing *Snhg9* expression. We tested this idea using a mouse model of CD11c^{+} cell depletion in which *Diphtheria* toxin receptor (DTR) is expressed from the *Cd11c* promoter (*34*). Selective depletion of CD11c^{+} cells by *Diphtheria* toxin administration increased *Snhg9* expression relative to controls (Fig. 5D). Because CD11c marks myeloid cells, including dendritic cells and macrophages, these results indicate that myeloid cells are required for microbial repression of intestinal *Snhg9* expression.

Bacteria regulate the expression of several key intestinal epithelial cell genes through an immune cell signaling relay involving primarily myeloid cells and group 3 innate lymphoid cells (ILC3s) (*2*, *35*–*37*). In this relay, bacteria activate myeloid cells through TLRs and MyD88, which then signal to ILC3 through the cytokine interleukin-23 (IL-23). Activated ILC3s then signal to intestinal epithelial cells through IL-22 (*2*, *35*–*37*). Having established a role for myeloid cells in microbial repression of *Snhg9* expression, we next tested for the involvement of ILCs. *Rag1*$^{-/-}$ mice, which lack T and B cells, had decreased expression of intestinal *Snhg9* compared with wild-type mice (Fig. 5E). This indicated that T and B cells are dispensable for the microbial repression of *Snhg9*, whereas decreased expression of *Snhg9* may be a result of the increased bacterial loads in the intestines of *Rag1*$^{-/-}$ mice (*36*). By contrast, depleting ILCs in *Rag1*$^{-/-}$ mice with the CD90.2 antibody (*38*) elevated *Snhg9* expression (Fig. 5F), which indicates a requirement for ILCs in microbial repression of *Snhg9* expression. Similarly, *Rag2*$^{-/-}$*;Il2rg*$^{-/-}$ mice, which lack immune cells (including ILCs) that are dependent on the IL-2 receptor γ chain, showed increased expression of intestinal *Snhg9* compared with *Rag1*$^{-/-}$ mice (Fig. 5G).

To further assess the requirement for ILC3, we analyzed *Snhg9* expression in *Rorc*$^{gfp/gfp}$ mice, which lack IL-22–producing cells, including ILC3 (*38*). *Rorc*$^{gfp/gfp}$ mice showed increased *Snhg9* expression compared with wild-type littermates, and antibiotic depletion of gut microbiota abolished this difference (Fig. 5H), which supports the idea that ILC3s relay microbial signals that repress *Snhg9* expression. Additionally, supplementing *Myd88*$^{-/-}$ mice with either IL-22 or IL-23 repressed *Snhg9* expression to the levels observed in conventional wild-type mice (Fig. 5I), consistent with the known involvement of these cytokines in the myeloid cell–ILC3 signaling circuit (*2*, *35*–*37*). Further, *Snhg9* expression was repressed by monocolonization of germ-free mice with bacterial species known to activate myeloid cell–ILC3 signaling (*2*, *35*, *37*), including *Salmonella enterica* Serovar Typhimurium (*S.* Typhimurium), a Gram-negative intestinal pathogen, and segmented filamentous bacteria (SFB), Gram-positive members of the intestinal microbiota (fig. S10, A and B). Thus, the microbiota suppresses *Snhg9* expression through a myeloid cell–ILC3 signaling relay.

Discussion and limitations

This study shows that the gut microbiota promotes lipid absorption and metabolism by repressing the expression of lncRNA *Snhg9*. This finding raises several additional questions for future study. First, are there specific components of the gut microbiota that promote lipid absorption through *Snhg9* repression? We found that *Snhg9* expression was selectively repressed by bacterial species that activate intestinal myeloid cell–ILC3 signaling, which provides a potential clue. However, more studies are needed to determine what bacterial components or characteristics enable activation of myeloid cell–ILC3 signaling and how microbial community composition affects *Snhg9* expression and lipid metabolism. Second, what is the evolutionary rationale for microbial regulation of lipid absorption? It is possible that intestinal lipid metabolism is linked to innate immune sensing of microbes to cope with an increased energy demand during colonization or infection, to provide lipid substrates or mediators that regulate intestinal immune cell development, or to enhance epithelial barrier function through reactive oxygen species production via lipid oxidation. Third, do these findings provide insight into the regulation of human lipid metabolism? *SNHG9* is conserved and expressed in humans (*19*), which raises the possibility that human *SNHG9* functions in a similar manner.

Conclusions

In this study, we show that *Snhg9* RNA regulates PPARγ activity by dissociating SIRT1 from CCAR2, providing insight into how a lncRNA regulates intestinal lipid metabolism. These findings advance our understanding of the complex epithelial cell networks that regulate lipid metabolism in response to microbial signals (*2*, *3*) (fig. S11). Ultimately, these results could suggest strategies for treating metabolic disease by targeting *Snhg9* and the microbiota.

REFERENCES AND NOTES

1. L. V. Hooper, T. Midtvedt, J. I. Gordon, *Annu. Rev. Nutr.* **22**, 283–307 (2002).
2. Y. Wang *et al.*, *Science* **357**, 912–916 (2017).
3. Z. Kuang *et al.*, *Science* **365**, 1428–1434 (2019).
4. K. A. Krautkramer, J. Fan, F. Bäckhed, *Nat. Rev. Microbiol.* **19**, 77–94 (2021).
5. V. Leone *et al.*, *Cell Host Microbe* **17**, 681–689 (2015).
6. N. Suárez-Zamorano *et al.*, *Nat. Med.* **21**, 1497–1501 (2015).
7. F. Bäckhed *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 15718–15723 (2004).
8. F. Bäckhed, J. K. Manchester, C. F. Semenkovich, J. I. Gordon, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 979–984 (2007).
9. S. Rabot *et al.*, *FASEB J.* **24**, 4948–4959 (2010).
10. V. K. Ridaura *et al.*, *Science* **341**, 1241214 (2013).
11. Y. Sanz, A. Santacruz, P. Gauffin, *Proc. Nutr. Soc.* **69**, 434–441 (2010).
12. A. Pascale *et al.*, *Endocrine* **61**, 357–371 (2018).
13. C. M. Hales, M. D. Carroll, C. D. Fryar, C. L. Ogden, *NCHS Data Brief* **360**, 1–8 (2020).
14. C. P. Ponting, P. L. Oliver, W. Reik, *Cell* **136**, 629–641 (2009).
15. A. Fatica, I. Bozzoni, *Nat. Rev. Genet.* **15**, 7–21 (2014).
16. A. Lin *et al.*, *Nat. Cell Biol.* **19**, 238–251 (2017).
17. D. Huang *et al.*, *Nat. Immunol.* **19**, 1112–1125 (2018).
18. T. R. Mercer, M. E. Dinger, J. S. Mattick, *Nat. Rev. Genet.* **10**, 155–159 (2009).
19. R.-H. Li *et al.*, *Cell Res.* **31**, 1088–1105 (2021).
20. D. K. L. Tsang *et al.*, *Gut Microbes* **14**, 2108281 (2022).
21. M. Kretz *et al.*, *Nature* **493**, 231–235 (2013).
22. S. Lee *et al.*, *Cell* **164**, 69–80 (2016).
23. F. Ferrè, A. Colantoni, M. Helmer-Citterich, *Brief. Bioinform.* **17**, 106–116 (2016).
24. M. Hamaguchi *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 13647–13652 (2002).
25. J.-E. Kim, J. Chen, Z. Lou, *Nature* **451**, 583–586 (2008).
26. W. Zhao *et al.*, *Nature* **451**, 587–590 (2008).
27. P. T. Pfluger, D. Herranz, S. Velasco-Miguel, M. Serrano, M. H. Tschöp, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 9793–9798 (2008).
28. A. S. Banks *et al.*, *Cell Metab.* **8**, 333–341 (2008).
29. P. Ferré, *Diabetes* **53**, S43–S50 (2004).
30. F. Picard *et al.*, *Nature* **429**, 771–776 (2004).
31. L. Qiang *et al.*, *Cell* **150**, 620–632 (2012).
32. K. Zebisch, V. Voigt, M. Wabitsch, M. Brandsch, *Anal. Biochem.* **425**, 88–90 (2012).
33. S. Vaishnava *et al.*, *Science* **334**, 255–258 (2011).
34. S. Jung *et al.*, *Immunity* **17**, 211–220 (2002).
35. T. Sano *et al.*, *Cell* **163**, 381–393 (2015).
36. J. M. Pickard *et al.*, *Nature* **514**, 638–641 (2014).
37. J. F. Brooks2nd *et al.*, *Cell* **184**, 4154–4167.e12 (2021).
38. G. F. Sonnenberg *et al.*, *Science* **336**, 1321–1325 (2012).

ACKNOWLEDGMENTS

We thank N. Salinas for assistance with mouse experiments, A. Lemoff (UT Southwestern Proteomics Core) for assistance with mass spectrometry, C. Liang (UT Southwestern Microarray Core) for assistance with RNA-seq, and A. Gewirtz (Georgia State University) for providing SFB-monocolonized mice. **Funding:** This work was supported by grant 2021YFA0805600 from the Ministry of Science and Technology of China (Y.W.), grants 32271213 (Y.W.) and 82188102 (T.L.) from the National Natural

Science Foundation of China, grant 2020R01006 from the Leading Innovative and Entrepreneur Team Introduction Program of Zhejiang (Y.W.), grant 226-2023-00134 from the Fundamental Research Funds for the Central Universities (Y.W.), NIH grants R01 DK070855 (L.V.H.) and R00 DK120897 (Z.K.), Welch Foundation grant I-1874 (L.V.H.), the Walter M. and Helen D. Bader Center for Research on Arthritis and Autoimmune Diseases (L.V.H.), and the Howard Hughes Medical Institute (L.V.H.). **Author contributions:** Y.W. and L.V.H. designed the research. Y.W., M.W., J.C., Y.L., C.D., P.R., G.Q., Z.H., T.S., B.H., K.R., and C.L.B. performed the research. T.L., X.D., and Z.S. provided experimental materials. Y.W., M.W., and Z.K. analyzed the data. Y.W. and L.V.H. wrote the paper. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** RNA-seq data and 16S ribosomal RNA gene sequencing data are available from the Gene Expression Omnibus (GEO) repository under accession no. GSE208020. All other data are available in the main text or the supplementary materials. **License information:** Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.ade0522
Materials and Methods
Figs. S1 to S11
Table S1
References (*39*–*49*)
MDAR Reproducibility Checklist

Submitted 21 July 2022; resubmitted 17 April 2023
Accepted 20 July 2023
10.1126/science.ade0522

CATALYSIS

Ternary NiMo-Bi liquid alloy catalyst for efficient hydrogen production from methane pyrolysis

Luning Chen[1]†, Zhigang Song[1,2]†, Shuchen Zhang[1]†, Chung-Kai Chang[3], Yu-Chun Chuang[3], Xinxing Peng[1], Chaochao Dun[4], Jeffrey J. Urban[4], Jinghua Guo[5], Jeng-Lung Chen[3]*, David Prendergast[4], Miquel Salmeron[1], Gabor A. Somorjai[1,6]*, Ji Su[1,7]*

Methane pyrolysis (MP) is a potential technology for CO_2-free hydrogen production that generates only solid carbon by-products. However, developing a highly efficient catalyst for stable methane pyrolysis at a moderate temperature has been challenging. We present a new and highly efficient catalyst created by modifying a Ni-Bi liquid alloy with the addition of Mo to produce a ternary NiMo-Bi liquid alloy catalyst (LAC). This catalyst exhibited a considerably low activation energy of 81.2 kilojoules per mole, which enabled MP at temperatures between 450 and 800 Celsius and a hydrogen generation efficiency of 4.05 ml per gram of nickel per minute. At 800 Celsius, the catalyst exhibited 100% H_2 selectivity and 120 hours of stability.

Hydrogen (H_2) is emerging as a promising clean energy solution on a global scale (*1*–*3*). However, most H_2 production (90 million tons in 2020) comes from fossil fuels, such as natural gas, oil, and coal, which result in substantial CO_2 emissions (about 900 million tons) (*4*). Water electrolysis is a green H_2 technology that uses renewable energy to produce CO_2-free H_2 (*5*). However, currently it only contributes 2% of H_2 production given its high cost (5 to 6 USD per kg_{H2}) and high energy consumption (286 kJ/mol_{H2}) (*6*). Methane (CH_4) pyrolysis (MP) is another CO_2-free method for H_2 production that in addition produces valuable carbon materials, such as graphene, carbon nanotubes, and fullerenes (*7*). Although the MP reaction requires only 37.5 kJ of energy to produce one mole of H_2, it still requires high reaction temperatures (>1000°C) to activate CH_4, leading to high energy demand, costly equipment, and unavoidable heat losses (*8*, *9*). Moderate reaction temperature would mitigate by-product formation (ethane, ethylene, acetylene, aromatics), minimizing H_2 separation and purification operations. Thus, it is essential to develop optimal catalysts that exhibit high catalytic activity, enabling a moderate operating temperature while also demonstrating excellent resistance to fouling and degradation.

Traditional supported transition metal catalysts (Ni, Co, Fe, Pt, or Pd) can catalyze MP under appropriate low reaction temperatures (500 to 600°C) with low apparent activation energy (*Ea*) values ranging from 65 to 96 kJ/mol, but they invariably suffer from deactivation by carbon coking and aromatics fouling (*10*, *11*). Molten liquid catalysts (MLCs) can overcome the deactivation issue by removing carbon products that float on top of the liquid catalysts and offer excellent durability. However, the high E_a of MLCs, ranging from 160 to 310 kJ/mol, demands high temperatures for CH_4 activation (Fig. 1A) (*12*–*14*).

Recent findings indicated that the Ni-Bi liquid alloy metal catalyst could catalyze MP. However, the strong interaction between the active metal sites and solvent metal created a cage with positively charged Bi atoms encapsulating negative charged Ni. This atomic arrangement obstructs the CH_4 reaction and results in low activity with a high activation energy E_a of 208 kJ/mol (*15*). Hence, reducing the cage effect of liquid metal catalysts is crucial to achieving a highly active catalyst for MP. Although liquid-metal catalysts have been used as new-generation catalysts in some specific catalytic processes (*16*–*18*), so far

[1]Chemical Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA. [2]John A. Paulson School of Engineering and Applied Sciences, Harvard University, Cambridge, MA 02138, USA. [3]National Synchrotron Radiation Research Center, Science-Based Industrial Park Hsinchu 300092, Taiwan. [4]Molecular Foundry, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA. [5]Advanced Light Source, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA. [6]Department of Chemistry, University of California-Berkeley, Berkeley, CA 94720, USA. [7]Energy Storage and Distributed Resources Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA.
***Corresponding author. Email: chen.jl@nsrrc.org.tw (J.-L.C.); somorjai@berkeley.edu (G.A.S.); jisu@lbl.gov (J.S.)**
†These authors contributed equally to this work.

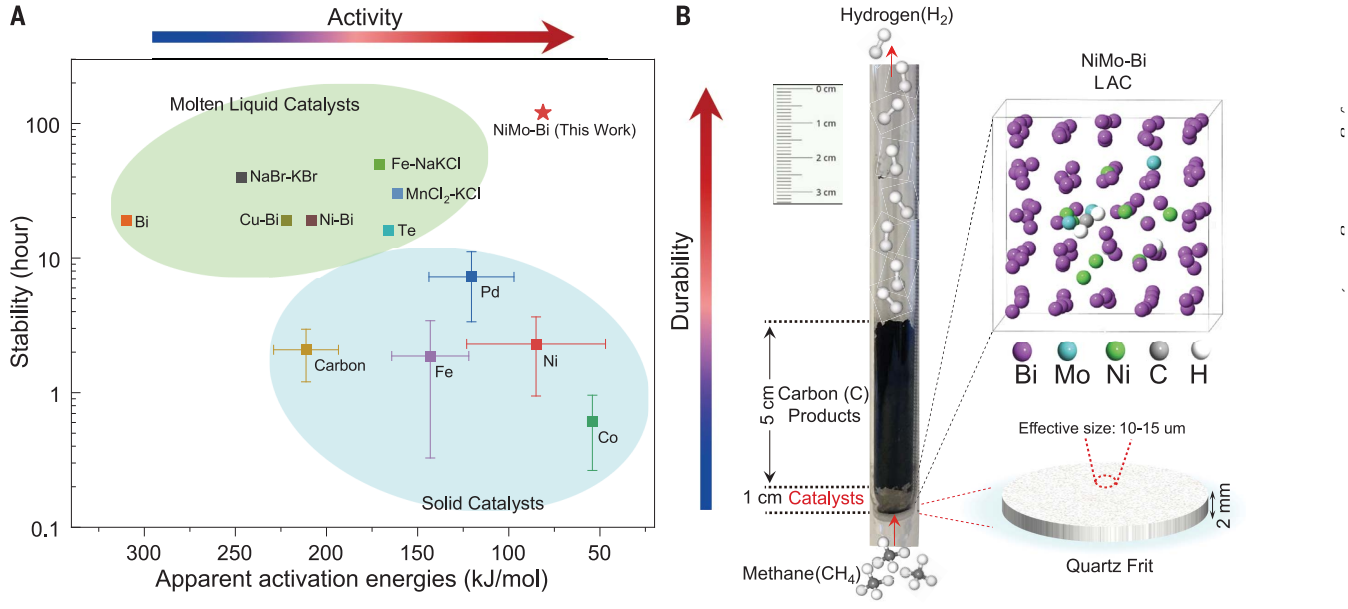


Fig. 1. Scheme of H_2 production from CH_4 pyrolysis. (**A**) Overview of catalysts for H_2 production from CH_4 pyrolysis, including the stability and activity of different solid and molten liquid catalysts. (**B**) Image of cooling down reactor after a long-time CH_4 pyrolysis reaction (detailed parameters and images in figs. S1 and S2).

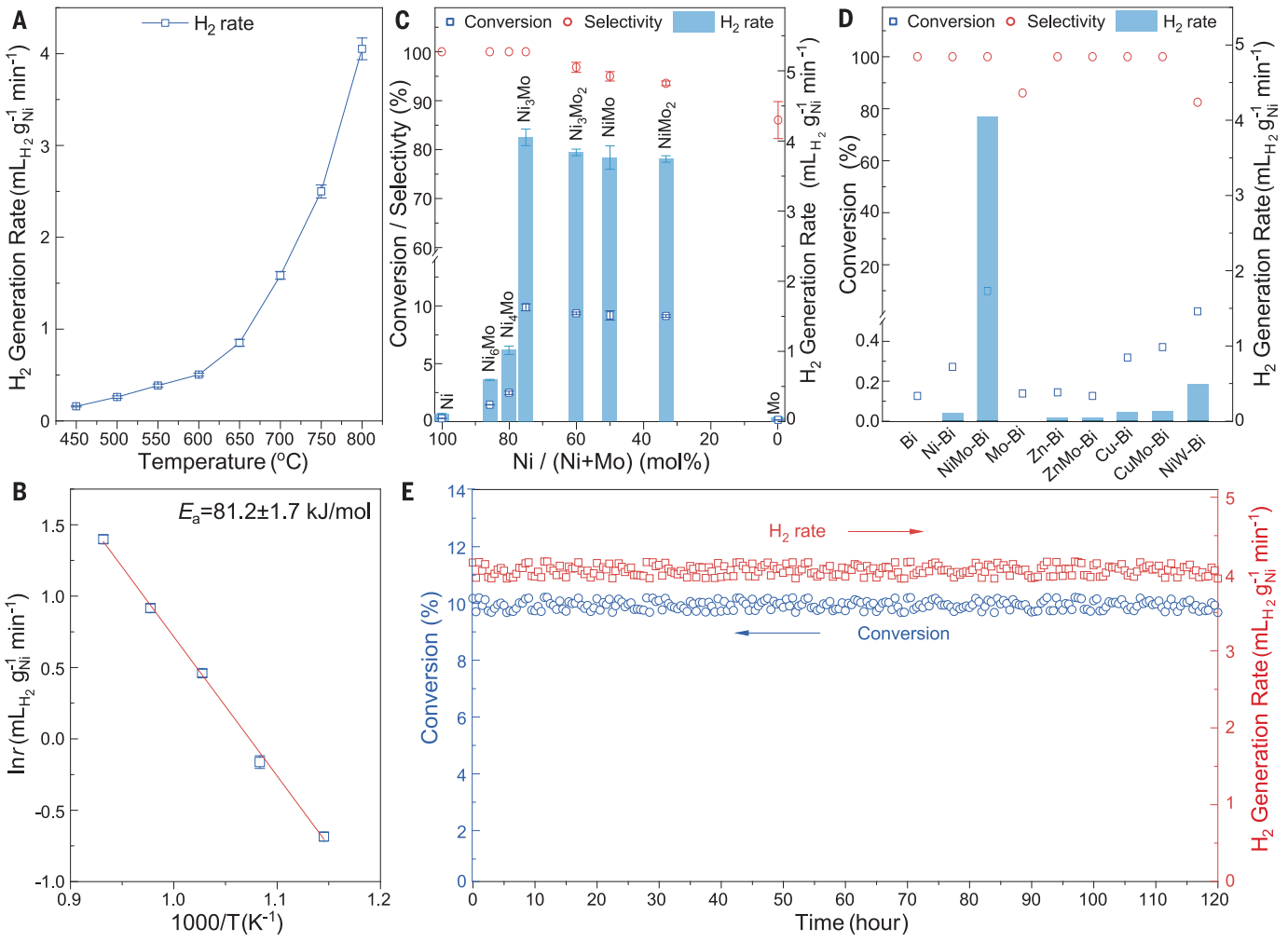


Fig. 2. Catalytic data. (**A**) H_2 generation rate under different reaction temperatures and (**B**) the apparent activation energy of CH_4 pyrolysis over NiMo-Bi liquid alloy catalyst (with a molar ratio of Ni to Mo of 3:1). Reaction conditions: 4 ml/min CH_4, pressure: 206 kPa (30 psi), catalyst height: <1 cm. (**C**) The conversion and selectivity (left) and H_2 generation rate (right) at 800°C over NiMo-Bi LAC with different composition. All catalysts contain the same Ni amount. (**D**) The conversion and selectivity (left) and H_2 generation rate (right) at 800°C over different liquid metal alloy catalysts. (**E**) The long-time stability measurement of CH_4 dehydrogenation over NiMo-Bi LAC (with a molar ratio of Ni to Mo of 3:1) at 800°C with 4 ml/min CH_4.

studies have not focused on tailoring the interaction between active sites and surrounding solution metal.

NiMo-Bi liquid alloy catalyst for MP

Compared with single-element metal materials, multi-element alloys have attracted attention in recent years because of their mechanical, physical, and chemical properties resulting from the entropy increase (*19*–*21*). The properties of alloy catalysts can be modified by additional elements, similar to how soluble metal complex catalysts can be modified by tuning ligands to modify active sites and their interaction with solvents (*22*–*24*). We added the third metal to regulate the interaction between the active metal and solution metals. We found that the introduction of Mo successfully decreases the cage effect from Bi atoms surrounding the Ni (*15*), as a result of the Ni-Mo interaction (Fig. 1B). The NiMo-Bi (Ni 2.3 wt %, Mo 1.3 wt %, and Bi 96.4 wt %) catalysts displayed the maximum MP activity with a high H_2 generation efficiency (4.05 $mL_{H2}\ g_{Ni}^{-1}\ min^{-1}$) at 800°C, which is 37 times faster than Ni-Bi catalysts. The E_a of our NiMo-Bi catalyst was 81.2 kJ/mol, substantially lower than other reported MLCs, approaching the level of supported solid metal catalysts (65 to 96 kJ/mol).

The preparation process of the NiMo-Bi catalyst was simple. For example, the best-performing NiMo-Bi catalyst (with an Ni-to-Mo molar ratio of 3:1) was prepared by directly reducing a mixture of nickel oxide, molybdenum oxide, and bismuth pellets in a quartz reactor with H_2 gas (figs. S1 and S2) (see details in supporting information). After the reduction, the mixture melted into the liquid metal solution under N_2 and was ready to be used for the MP reaction. We used high-resolution powder x-ray diffraction (XRD) to identify the structure of the liquid NiMo-Bi catalyst after cooling down to room temperature under an inert atmosphere. We identified a major crystal Bi phase with a symmetry group of *R-3m* (ICSD-64703), and a minor Bi_3Ni phase with *Pnma* symmetry (ICSD-391336) according to the fitting results (fig. S3) (*25*). No appreciable amounts of phases related to the Mo component were detected compared with the Ni-Bi sample. We found that the Ni-Mo interaction could enhance the solubility of Mo in Bi, and more evidence will be provided in the spectroscopy and theoretical study sections. We next studied the phase transitions in a melting process by in situ high-temperature XRD (fig. S4). The crystal phase of Bi disappeared at 260°C near the Bi melting point of 270°C. The Bi_3Ni phase disappeared at 420°C (*26*), indicating that above this temperature, the NiMo-Bi catalyst remained in the liquid state, with Ni and Mo homogeneously distributed in the liquid solution (fig. S5). In situ high-temperature energy-dispersive spectroscopy mapping showed that

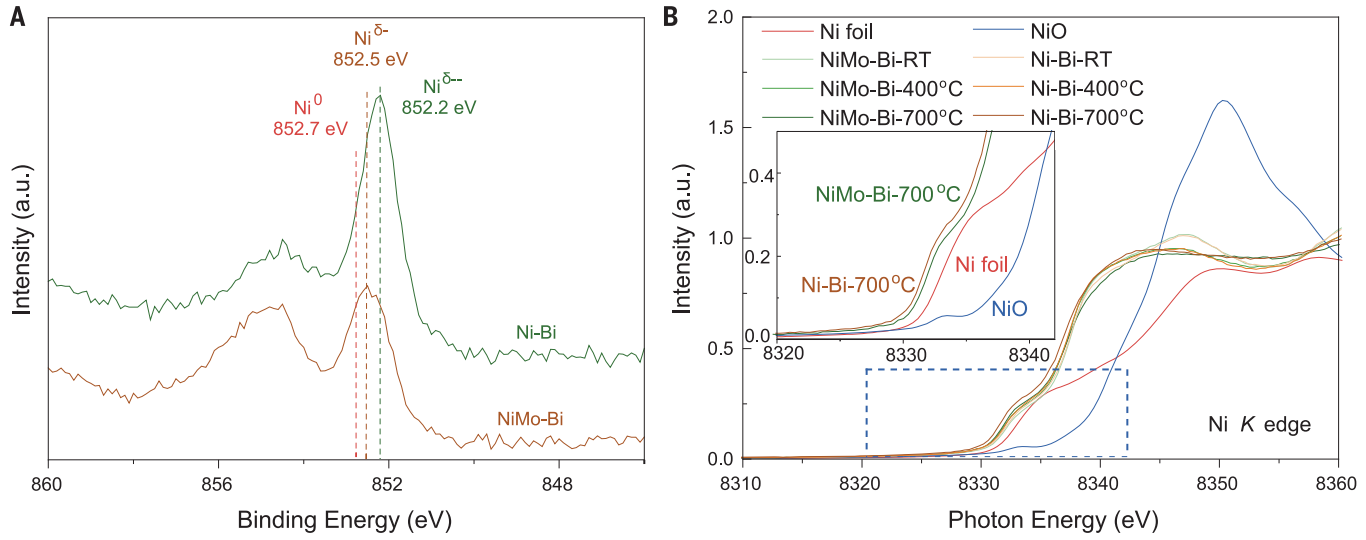


Fig. 3. Statement of Ni species of Ni-Bi and NiMo-Bi liquid alloy catalysts. (**A**) Nickel 2p x-ray photoelectron spectroscopy (XPS) of Ni-Bi and NiMo-Bi catalysts at room temperature. (**B**) Ni K-edge x-ray absorption near edge spectroscopy (XANES) spectra of Ni-Bi and NiMo-Bi catalysts at room temperature (RT) and operating temperature.

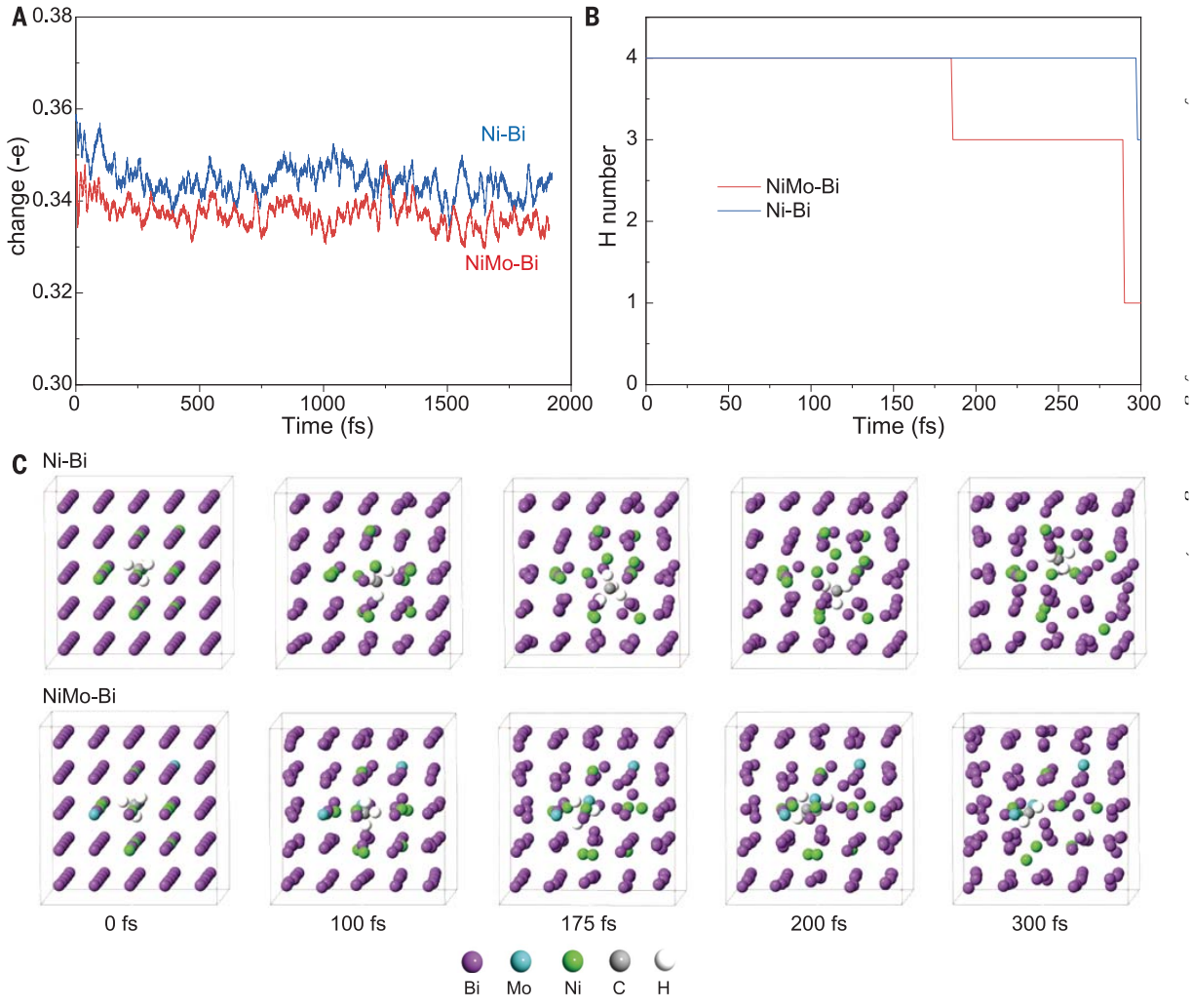


Fig. 4. Molecular dynamics simulation. (**A**) Average transferred electrons to a Ni atom. Molecular dynamics is performed using alloys of Ni-Bi and NiMo-Bi at 1500 K. (**B**) H atomic number around a C atom as a function of time. Molecular dynamics is performed using a CH_4 molecule in the alloy of NiMo-Bi and Ni-Bi at 1500 K. (**C**) Snaps of CH_4 in the liquid alloy of Ni-Bi and NiMo-Bi (detailed process in movies S5 and S6).

at 500°C, the Ni, Mo, and Bi were uniformly distributed in the liquid alloy (fig. S6).

Catalytic studies

After catalyst preparation, CH_4 was introduced in the reactor and passed through the liquid catalyst to investigate the MP performance at different temperatures with a flow rate of 4 ml/min CH_4 and a 206 kPa (30 psi) pressure. The residence time of CH_4 through the liquid alloy catalyst was around 0.13 min.

No H_2 was detected below 400°C. At 450°C, the catalysts were completely melted (fig. S5), and H_2 was detected with a generation rate of 0.16 $mL_{H2}\ g_{Ni}^{-1}\ min^{-1}$. High temperature favored H_2 production and CH_4 conversion. Increasing the temperature to 800°C, the H_2 generation rate increased to 4.05 $mL_{H2}\ g_{Ni}^{-1}\ min^{-1}$ and the CH_4 conversion reached 9.87% (Fig. 2A), which is 37 times higher than the Ni-Bi catalyst (0.11 $mL_{H2}\ g_{Ni}^{-1}\ min^{-1}$). No other by-products were detected under the operation temperatures (fig. S7), indicating complete CH_4 decomposition. Notably, this reaction system avoided the production of aromatics that result in fouling by acetylene by-products, which would cause catalyst deactivation and reaction blocking (*27*, *28*). The E_a of 81.2 kJ/mol was determined from the Arrhenius plots (Fig. 2B). This value is much lower than in all reported MLCs (166 to 310 kJ/mol) and comparable to that of solid metal catalysts (*13*).

The catalytic MP performance was further investigated under different reaction conditions (fig. S8). Lower flow rates or CH_4 concentration led to an increase in CH_4 conversion. When the CH_4 flow rate decreased to 1 ml/min, the conversion of CH_4 increased to 11.2%. Increasing the CH_4 pressure to 310 kPa (45 psi), the CH_4 conversion increased to 13.1% as a result of CH_4 solubility enhancement. Notably, the residence time of CH_4 could be tuned by the catalyst height. Longer residence times increase the CH_4 conversion. For a 5-cm-tall NiMo-Bi catalyst column, CH_4 conversion raised to 16.3%, which is higher compared with previous Ni-Bi catalysts (13%) even with higher catalyst columns (8 cm) and higher temperature (1000°C) (*15*). When 10% CH_4 diluted in N_2 was fed into the 5-cm-tall NiMo-Bi catalyst, the CH_4 conversion could reach up to 51.7% under 800°C and 2 bar of total pressure. These results indicated that CH_4 conversion could be further enhanced by reactor engineering optimization. Considering the considerably higher activity of our catalyst and similar catalyst physical properties to the Ni-Bi catalyst (*15*), it is reasonable to assume that achieving a conversion close to the reaction equilibrium is feasible.

To determine the effect of different elements on catalytic results, we investigated the performance of liquid alloy catalysts with different Ni-Mo-Bi ratios (Fig. 2C and S9). Pure liquid Bi displayed no measurable activity for MP below 800°C. The Ni-Bi catalyst (without Mo) generated H_2 at a rate of 0.02 $mL_{H2}\ g_{Ni}^{-1}\ min^{-1}$ starting at 650°C, which increased to 0.11 $mL_{H2}\ g_{Ni}^{-1}\ min^{-1}$ at 800°C. Introducing Mo increased the reaction activity and reduced the onset temperature for MP. The H_2 generation rate over the Ni_6Mo-Bi catalyst was 0.59 $mL_{H2}\ g_{Ni}^{-1}\ min^{-1}$ at 800°C and reached 1.02 $mL_{H2}\ g_{Ni}^{-1}\ min^{-1}$ for Ni_4Mo-Bi. The best catalytic performance was achieved at a ratio of Ni to Mo of 3:1. As the concentration of Mo was further increased, the H_2 generation rate slightly decreased, whereas the H_2 selectivity also decreased because of the formation of hydrocarbon by-products. These results demonstrate that in the NiMo-Bi liquid alloy system Ni was the active metal, with Bi acting as the solvent. Further, Mo, which has strong interaction with Ni, performed as the regulatory metal and enhanced the reaction activity by modulating the interaction of the solvent Bi metal and the active Ni metal.

We further investigated liquid alloy catalyst systems of different compositions, all in the liquid phase under the reaction temperature (Fig. 2D and fig. S10). No catalytic activity was detected for either Zn-Bi or ZnMo-Bi catalysts. The Cu-Bi catalyst had similar catalytic activity to Ni-Bi (*29*). However, in contrast to the NiMo-Bi system, the addition of Mo did not appreciably enhance the reaction activity of the CuMo-Bi catalyst. Furthermore, when W was used instead of Mo in the NiMo-Bi system, despite its similar chemical and physical properties, the enhancement observed was much smaller compared with Mo, especially at high temperatures, due to the weaker interaction between Ni and W compared to that between Ni and Mo (*30*).

In addition to high activity and selectivity, the NiMo-Bi catalyst also exhibited excellent stability in the MP reaction. After 120 hours of measurement at 800°C with a methane flow rate of 4 ml/min, the NiMo-Bi maintained excellent activity and selectivity without any deactivation relative to the initial performance (Fig. 2E), which is better than most MP catalysts including MLCs and traditional solid catalysts (Fig. 1A). The Ni and Mo were still uniformly dispersed in the Bi solution without any aggregation, suggesting that the NiMo-Bi could maintain stability over an even longer period. During the reaction, the carbon product was segregated from the catalysts and accumulated on the catalyst surface (fig. S11). No sticky aromatic fouling species were detected on the wall of the quartz reactor. By further washing the carbon products with acetone, no aromatic products were detected by GC-MS. Notably, although the carbon could also catalyze the MP reaction, under our reaction condition the activity of the NiMo-Bi catalyst was much higher than that of the formed carbon, indicating that all MP activity comes from NiMo-Bi catalysts. To further confirm this, 100-mg carbon products were used as a catalyst, which displayed low activity (0.6% conversion) and fast deactivation rate (36% in 5 min) under a similar reaction condition (fig. S12).

Spectroscopy studies

X-ray photoelectron spectroscopy (XPS) and in situ x-ray absorption near-edge spectroscopy (XANES) were applied to further investigate the chemical state of the species in the liquid alloy catalysts. XPS (Fig. 3A) results show that the binding energy of the Ni 2p level in the Ni-Bi catalyst (852.2 eV) was 0.5 eV lower than that of metallic Ni (852.7 eV) (*31*), indicating that Ni was negatively charged by electron transfer from surrounding Bi, in agreement with previous work (*32*). After Mo introduction, the peaks of Ni in NiMo-Bi shifted to a higher binding energy (852.5 eV), suggesting that the Ni in NiMo-Bi catalyst is closer to the metallic state but still negatively charged.

In situ high-temperature XANES (fig. S13) was used to explore the Ni electronic structure at the reaction temperature (Fig. 3B and fig. S14). Consistent with XPS results, the adsorption edge energies of Ni in Ni-Bi and NiMo-Bi LACs were lower than that of the Ni reference foil at room temperature, indicating its negative charge resulting from the interaction with Bi (*33*). Heating up catalysts under an inert atmosphere, the adsorption edge energies of both Ni-Bi and NiMo-Bi moved to lower values, implying that in the liquid state, Ni is well mixed with Bi, favoring electron transfer from Bi to Ni (*15*, *34*). After Mo introduction, the strong interaction between Ni and Mo modulated the electronic state of Ni and reduced the interaction between Ni and Bi (*35*). Additionally, the XANES spectra of NiMo-Bi at Mo *K*-edge showed that at both high temperature and room temperature the valence state of Mo is different from that in the metallic foil and molybdenum oxide, which indicated that there is no Mo metal or molybdenum oxide in the liquid alloy. Further, no Mo-Mo or Ni-Ni bonds were detected in the NiMo-Bi liquid alloy according to room temperature extended x-ray absorption fine structure spectra (fig. S15). The different radial distances and coordination environment of Ni and Mo of NiMo-Bi catalysts, compared with metal foil counterparts, further proved that Ni and Mo were uniformly dispersed without any aggregation (*36*). Moreover, the difference between the Ni-Bi and NiMo-Bi catalysts indicated the interaction between Ni and Mo, which is also different from the traditional Ni-Bi alloy. These spectroscopy studies successfully verified the existence of interactions between Ni and Mo in the NiMo-Bi melt, which resulted in Ni being less negatively charged and the enhancement of Mo solubility in the Ni-Bi system.

Theoretical studies

A theoretical molecular dynamics simulation based on density functional theory was conducted

to gain deeper insights into the electronic structure and reaction process in the liquid alloy catalysts (the details of the calculation are shown in the supplementary information). First, we simulated the dissolution process of the Mo_2 dimer and Mo cluster (fig. S16 and movies S1 and S2). In the pure molten Bi, the distance of two Mo atoms of Mo_2 dimer did not change as a result of the limited solubility of Mo (*37*), but in the presence of Ni the Mo-Mo distance would increase. Similar results were found in the Mo cluster in Bi-Ni liquid alloy. These results demonstrate that the solubility of Mo could be enhanced as a result of Ni-Mo interaction leading to Mo uniformly dispersing in the liquid alloy instead of aggregating. Then we monitored the average Ni electron charge in Ni-Bi and NiMo-Bi catalysts (Fig. 4A). In accordance with XPS and XANES results, the Ni dissolved in Bi with or without Mo are both negatively charged because of the difference of electron affinity between Ni and Bi. After Mo introduction the charge of the Ni atoms decreased, which could potentially increase the Ni mobility as well as decrease the interaction between Ni and surrounding Bi atoms (*38*).

To further understand this phenomenon, we embedded a Ni atom in the Bi solution and carried out a molecular dynamics simulation at 1500 K. The Ni atom was surrounded by Bi atoms (fig. S17 and movie S3). However, after Mo introduction, the Ni atom broke through the surrounding Bi atoms, making it more accessible to CH_4 molecules as catalytic active sites (fig. S18 and movie S4). Subsequently, we introduced CH_4 in the NiMo-Bi and Ni-Bi systems to follow the dissociation process at 1500 K (Fig. 4C and movies S5 and S6). The number of H of CH_4 was calculated and the results are listed in Fig. 4B. The strong interaction between Ni and Bi made it more difficult for the CH_4 molecules to reach the caged Ni active sites, resulting in a reaction time to dissociate of about 300 femtosecond (fs) for the Ni-Bi system. However, the introduction of Mo weakened this cage effect and increased the possibility of interaction between Ni and CH_4, which caused CH_4 to start to dissociate at about 180 fs for the NiMo-Bi system. If we introduced W to replace Mo, the CH_4 started to dissociate at 252 fs for NiW-Bi, due to the weak interaction between Ni and W at a high temperature (fig. S19 and movie S7). All these results are in agreement with our experimental observation.

Discussion

The MP process is widely recognized as a promising approach for clean H_2 production. Nevertheless, it is suffering from a lack of a catalyst capable of efficiently, selectively, and stably catalyzing the MP process under mild reaction temperatures. We developed a NiMo-Bi liquid alloy catalyst that could simultaneously perform high efficiency, selectivity, and durability for methane pyrolysis at mild temperatures.

It is worth noting that the NiMo-Bi liquid alloy catalyst also exhibited high activity for the pyrolysis of other natural gas components, such as ethane and propane (fig. S20). This indicates that this liquid metal alloy catalyst could be further applied in H_2 production from other sources, such as biomass and plastic (*39*, *40*). Further, more-efficient soluble multi-elementary liquid alloy catalysts with different compositions are also under development, which may break through the recent reaction limitations and change the future of catalysis.

REFERENCES AND NOTES

1. L. P. Bicelli, *Int. J. Hydrogen Energy* **11**, 555–562 (1986).
2. "U.S. National Clean Hydrogen Strategy and Roadmap" (United States Department of Energy Press, 2023); https://www.hydrogen.energy.gov/pdfs/us-national-clean-hydrogen-strategy-roadmap.pdf.
3. J. O. Abe, A. P. I. Popoola, E. Ajenifuja, O. M. Popoola, *Int. J. Hydrogen Energy* **44**, 15072–15086 (2019).
4. "HYDROGEN STRATEGY Enabling A Low-Carbon Economy" (Office of Fossil Energy, United States Department of Energy Press, 2020) https://www.energy.gov/sites/prod/files/2020/07/f76/USDOE_FE_Hydrogen_Strategy_July2020.pdf.
5. A. J. Shih *et al*., *Nat. Rev. Methods Primers* **2**, 84 (2022).
6. O. Schmidt *et al*., *Int. J. Hydrogen Energy* **42**, 30470–30492 (2017).
7. N. Muradov, *Int. J. Hydrogen Energy* **26**, 1165–1175 (2001).
8. C. Guéret, M. Daroux, F. Billaud, *Chem. Eng. Sci.* **52**, 815–827 (1997).
9. H. F. Abbas, W. W. Daud, *Int. J. Hydrogen Energy* **35**, 1160–1190 (2010).
10. N. S. N. Hasnan *et al*., *Mater. Renew. Sustain. Energy* **9**, 8 (2020).
11. J. X. Qian *et al*., *Int. J. Hydrogen Energy* **45**, 15721–15743 (2020).
12. B. Parkinson, J. W. Matthews, T. B. McConnaughy, D. C. Upham, E. W. McFarland, *Chem. Eng. Technol.* **40**, 1022–1030 (2017).
13. M. Msheik, S. Rodat, S. Abanades, *Energies* **14**, 3107 (2021).
14. K. Wang, W. S. Li, X. P. Zhou, *J. Mol. Catal. Chem.* **283**, 153–157 (2008).
15. D. C. Upham *et al*., *Science* **358**, 917–921 (2017).
16. M. A. Rahim *et al*., *Nat. Chem.* **14**, 935–941 (2022).
17. N. Taccardi *et al*., *Nat. Chem.* **9**, 862–867 (2017).
18. K. Kalantar-Zadeh, M. A. Rahim, J. Tang, *Acc. Mater. Res.* **2**, 577–580 (2021).
19. P. C. Chen *et al*., *Science* **352**, 1565–1569 (2016).
20. E. P. George, D. Raabe, R. O. Ritchie, *Nat. Rev. Mater.* **4**, 515–534 (2019).
21. K. Kusada, M. Mukoyoshi, D. Wu, H. Kitagawa, *Angew. Chem. Int. Ed.* **61**, e202209616 (2022).
22. X. Chen *et al*., *Nature* **592**, 712–716 (2021).
23. Q. Pan *et al*., *Science* **374**, 984–989 (2021).
24. W. A. Herrmann, C. W. Kohlpaintner, *Angew. Chem. Int. Ed.* **32**, 1524–1544 (1993).
25. J. Kumar *et al*., *Supercond. Sci. Technol.* **24**, 085002 (2011).
26. R. Sheikhi, J. Cho, *J. Mater. Sci. Mater. Electron.* **29**, 19034–19042 (2018).
27. A. Jess, *Fuel* **75**, 1441–1448 (1996).
28. X. Guo *et al*., *Science* **344**, 616–619 (2014).
29. C. Palmer *et al*., *ACS Catal.* **9**, 8337–8345 (2019).
30. H. Okamoto, P. R. Subramanian, L. Kacprzak, Binary Alloy Phase Diagrams, Second Edition (ASM International, Materials Park, OH, 1990).
31. P. L. Tam, L. Nyborg, *Surf. Coat. Tech.* **203**, 2886–2890 (2009).
32. H. B. Michaelson, *J. Appl. Phys.* **48**, 4729–4733 (1977).
33. B. Das *et al*., *Phys. Chem. Chem. Phys.* **24**, 4415–4424 (2022).
34. H. Liu *et al*., *Nat. Catal.* **4**, 202–211 (2021).
35. S. Guo, C. T. Liu, *Prog. Nat. Sci.* **21**, 433–446 (2011).
36. Y. H. Liu *et al*., *Sci. Adv.* **9**, eadf9931 (2023).
37. A. V. Abramov *et al*., *Metals (Basel)* **13**, 366–375 (2023).
38. R. Chen *et al*., *Acta Mater.* **144**, 129–137 (2018).
39. H. Balat, E. Kırtay, *Int. J. Hydrogen Energy* **35**, 7416–7426 (2010).
40. X. Jie *et al*., *Nat. Catal.* **3**, 902–912 (2020).

ACKNOWLEDGMENTS

Funding: This work was primarily supported by the director, Office of Basic Energy Sciences, Division of Chemical Sciences, Geological and Biosciences of the US Department of Energy under contract DE-AC02-05CH11231. This work was also partially supported by the Hydrogen Materials Advanced Research Consortium (HyMARC), established as part of the Energy Materials Network by the US Department of Energy (DOE), Office of Energy Efficiency and Renewable Energy, Fuel Cell Technologies Office, under the same contract number, DE-AC02-05CH11231. The x-ray absorption near edge structure (XANES) work was supported by the Ministry of Science and Technology of Taiwan under grant MOST 111-2112-M-213-016 and NSTC 112-2112-M-213-016. Work by L.C., S.Z., C.D., J.J.U., D.P., and J.S. was performed as part of a User Project at The Molecular Foundry of the Lawrence Berkeley National Laboratory, and J.G. received support from the research resource of The Advanced Light Source, which is supported by the Office of Science of the US Department of Energy, also under the same contract (DE-AC02-05CH11231). For computational resources, we thank National Energy Research Scientific Computing Center (NERSC) and High-Performance Computing Services (HPCS) at Lawrence Berkeley National Laboratory. **Author contributions:** L.C., S.Z., G.A.S., and J.S. conceived of the project directions. L.C., S.Z., and J.S. designed the experiments. L.C., Z.S., S.Z., C.-K.C., Y.-C.C., X.P., C.D., and J.-L.C. performed the experiments. L.C. and Z.S. wrote the manuscript. L.C., S.Z., J.J.U., J.G., J.-L.C., M.S., G.A.S., and J.S. reviewed the manuscript, which was approved by all authors. **Competing interests:** L.C. and J.S. are inventors on a US provisional patent application for the process described herein. The remaining authors declare no competing interests. **Data and materials availability:** All data are available in the main text or the supplementary materials. **License information:** Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.sciencemag.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adh8872
Materials and Methods
Figs. S1 to S21
References (*41*–*44*)
Movies S1 to S7

Submitted 20 March 2023; accepted 24 July 2023
10.1126/science.adh8872

MATERIALS SCIENCE

Pneumatic cells toward absolute Gaussian morphing

Tian Gao, José Bico, Benoît Roman*

On a flat map of the Earth, continents are inevitably distorted. Reciprocally, curving a plate simultaneously in two directions requires a modification of in-plane distances, as Gauss stated in his seminal theorem. Although emerging architectured materials with programmed in-plane distortions are capable of such shape morphing, an additional control of local bending is required to precisely set the final shape of the resulting three-dimensional surface. Inspired by bulliform cells in leaves of monocotyledon plants, we show how the internal structure of flat panels can be designed to program bending and in-plane distortions simultaneously when pressurized, leading to a targeted shell shape. These surfaces with controlled stiffness and fast actuation are manufactured using consumer-grade materials and open a route to large-scale shape-morphing robotics applications.

Plants leaves and petals provide a good example of shape morphing induced by differential growth and are a source of inspiration for engineering (*1*). Materials capable of changing shape find numerous applications ranging from the fabrication of complex microstructures (*2*) and the manipulation of fragile objects (*3*, *4*) or soft-robotics devices (*5*–*8*) to locomotion in complex environments (*9*) or the design of deployable shelters (*10*). Although basic grabbing devices rely on bending linear beams through bilayer effects (*11*), shape morphing of surfaces poses a much greater geometrical challenge. If distances along the plane of a surface (i.e., the metrics) are conserved, then its Gaussian curvature (i.e., the product of its two principal curvatures) is strictly conserved, restricting the achievable shapes to families of isometries (e.g., cylinders or cones for a sheet of paper). General morphing of a surface, such as when a plane takes on a doubly curved shape, referred to here as Gaussian morphing, requires metric distortion (*12*). In natural structures, such metric changes result from differential growth (*13*, *14*). Engineered systems rely on equivalent nonhomogeneous transformations such as swelling of hydrogels (*15*–*18*), relaxation of liquid crystal elastomers (*19*) or prestretched polymeric filaments (*20*, *21*), inflatable structures (*22*–*24*), and origami or kirigami tessellation (*25*–*27*). Although the shape of a surface is rigorously defined by both curvature and metric tensors, these two parameters are generally not addressed simultaneously except for a few elegant exceptions involving very compliant materials relying on relatively slow swelling actuation (*28*–*30*). Programming complete morphing in stiff structures with fast actuation that are relevant for engineering applications remains a challenge.

Laboratoire de Physique et Mécanique des Milieux Hétérogènes (PMMH), CNRS, ESPCI Paris, Université PSL, Sorbonne Université, Université Paris Cité, 75005 Paris, France.
*Corresponding author. Email: benoit.roman@espci.fr.

In the realm of botany, the leaves of monocotyledon plants such as corn are able to curl inward reversibly under dry conditions, limiting evaporation (Fig. 1A). Actuation relies on peculiar bulliform cells that change their volume depending on turgor pressure (*31*, *32*), producing in-plane extension or contraction of one side of the surface leaf. Inspired by example from nature, we designed thin panels embedded with inflatable cells referred to as pneumatic Gaussian cells. Both in-plane contraction and angular deflection of the panels can be programmed simultaneously through the cell's design, leading to stiff, three-dimensional (3D) structures. This qualitative step constitutes an important milestone toward versatile morphing robotics applications.

Contracting and bending simultaneously

The elementary unit of a Gaussian cell is based on trapezoidal channels 3D printed on a layer of airtight fabric (Fig. 1B). A heat-sealed layer of the same fabric closes the cell (for the detailed fabrication process, see the supplementary materials and movie S1). The different channels embedded in the thin sandwich are connected to a pressure source. Pneumatic actuation of the cells induces the transformation of the initially flat panel into a complex shape such as a curve-folded origami structure (Fig. 1C and movie S1). Figure 2A presents the deployment

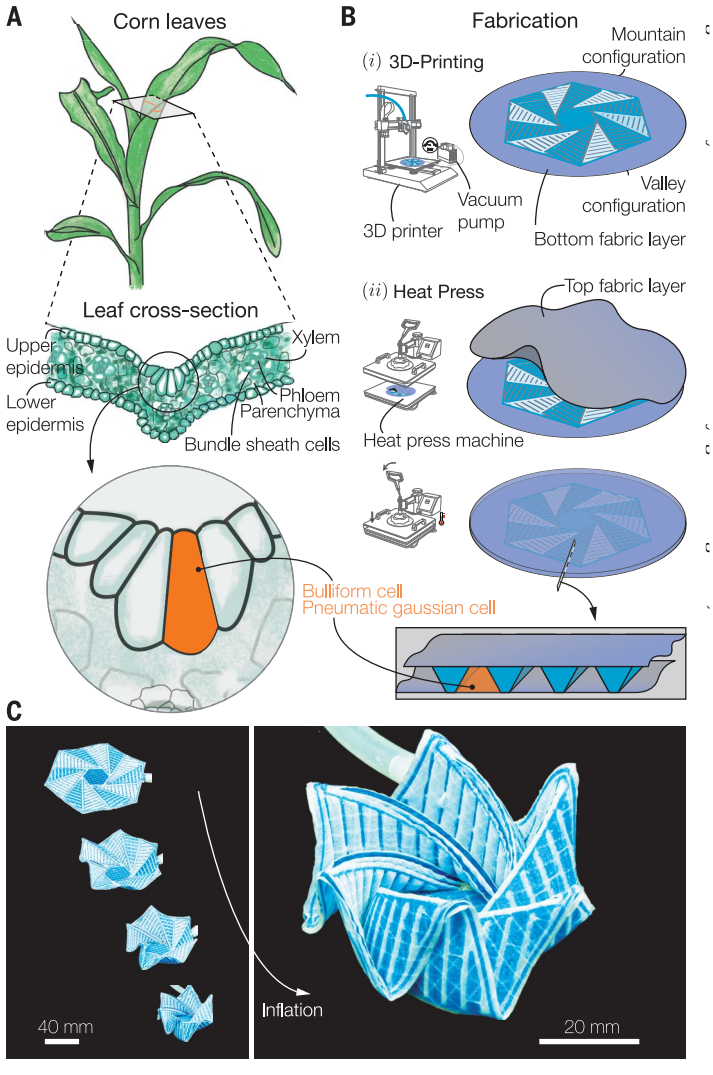


Fig. 1. Production of pneumatic Gaussian cells inspired by leaf bulliform cells. (**A**) Leaves of corn open or close as a function of weather conditions by inflating of deflating bulliform cells under turgor pressure. (**B**) Fabrication of a shape-morphing structure based on pneumatic Gaussian cells. (i) Tilted thermoplastic polyurethane (TPU) walls are 3D printed over a first layer of TPU-coated fabric. (ii) Gaussian cells (orange) are formed by sealing a second layer of fabric on the top of the structure with a heating press. (**C**) Upon inflation (see movie S1), an initially flat panel composed of programmed Gaussian cells self-shapes into a complex 3D structure inspired by the starshade project (*42*, *43*).

mechanism of a plate composed of juxtaposed parallel Gaussian cells with spacial periodicity ℓ. Because the base and the top of the cells are sealed with highly bendable but inextensible fabric, the cells tend to bulge into circular arcs upon inflation, inducing an inward displacement and rotation of the lateral walls. The general case of a symmetric trapezoidal cross-section (orange cells in Fig. 2A) is characterized by the thickness of the plate H, the shortest width W, and the internal angle β of the cell walls (see the supplementary materials for the case of asymmetric cells). The inflated shape of a Gaussian cell is readily obtained from energy minimization (for details, see the supplementary materials). The resulting folding angle γ of the cell is a solution of

$$\left(\beta-\frac{\gamma}{2}\right) = \sin\beta\cos\left[\left(\beta-\frac{\gamma}{2}\right)\left(\frac{W}{H}\cot\beta+1\right)\right] \quad (1)$$

Inflation also reduces the length of the mid-segment of the cell by a factor λ (Fig. 2B). This quantity referred to as in-plane contraction is solution of

$$\lambda = \frac{\cos\left(\beta-\frac{\gamma}{2}\right)}{\left(\beta-\frac{\gamma}{2}\right)\cos\frac{\gamma}{2}}\sin\left[\left(\beta-\frac{\gamma}{2}\right)\left(\frac{W}{H}\cot\beta+1\right)\right]\left(\frac{W}{H}\cot\beta+1\right) \quad (2)$$

The two extreme cases highlighted in yellow and green in Fig. 2 correspond to cells of rectangular ($\beta = 0°$) and triangular ($W/H = 0$) cross-sections that only provide in-plane contraction or deflection, respectively. Experimental values of γ and λ obtained over a wide range of geometrical parameters show excellent agreement (Fig. 2, B and C) with theoretical predictions (Eqs. 1 and 2). Figure 2, C and D, presents the wide palette of contraction ratio and folding angle as a function of the geometry of inflatable cells parameterized by their internal angle β and aspect ratio W/H. Large folding angles are found in the upper part of the parameter space, whereas large internal angles β promote asymmetry across the thickness of the cell, and values as large as $\gamma > 90°$ can be obtained. The largest contraction is intuitively obtained in the lower right corner, for large W/H. Contraction ratio and folding angle vary along different patterns in the parameter space. For instance, point (ii) in Fig. 2, D and E, corresponds to the same contraction as point (i) but to a larger folding angle. An important consequence is that arbitrary combinations of (λ,γ) can be programmed.

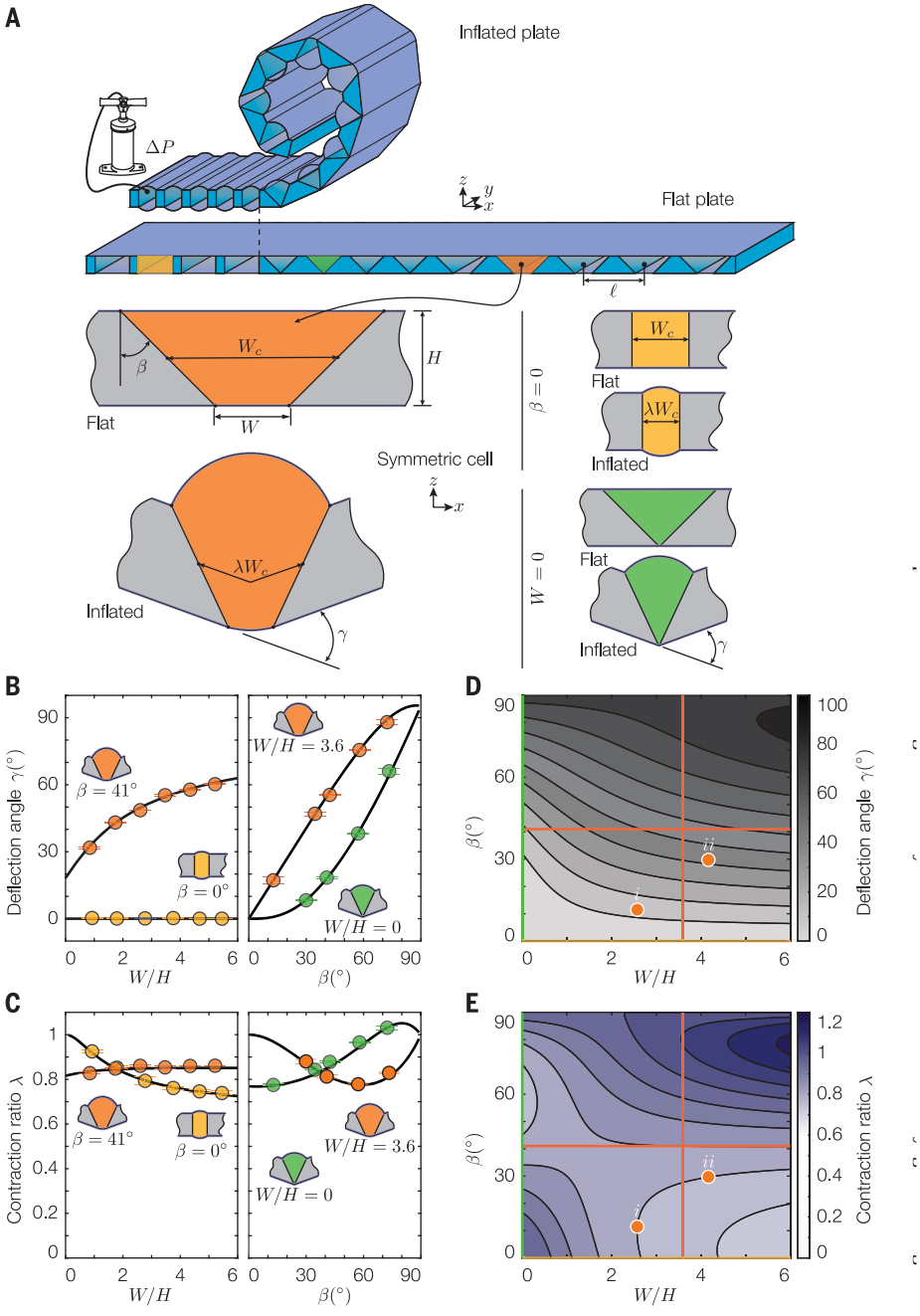


Fig. 2. Programming deformation of Gaussian cells. (**A**) Schematic illustration of the deployment of a flat plate based on three types of symmetric cells. Shown is a generic trapezoidal cell (orange) with walls of thickness H tilted by an angle β and separated by a distance W. Upon inflation, the cell induces a local deflection γ and a contraction λ with respect to the median line W_c. Rectangular cell (yellow) and triangular cell (green) configurations correspond to vertical walls ($\beta = 0$) or to contacting bases ($W = 0$), respectively. (**B** and **C**) Deflection angle γ and contraction ratio λ as a function of the dimensionless width W/H and of the tilt angle β for fixed values of β or W/H, respectively. Solid black lines correspond to theoretical predictions (Eqs. 1 and 2). (**D** and **E**) Landscape of programmable deformations. Shown are theoretical contour maps of the contraction ratio λ and deployed angle γ in a general configuration.

Stiff pneumatic Gaussian cells

Gaussian morphing requires large in-plane deformations and is therefore usually associated with low stiffness (*15*, *17*, *19*, *28*). Pneumatic Gaussian cells provide substantial load-bearing capabilities, which can be adjusted not only by pressure but also by the geometry of the cell, a feature distinct from traditional inflatable structures. Their stretching stiffness is assessed by testing the mechanical response of a symmetric cell ($\beta = 0°$) loaded in the direction perpendicular to the channel walls (Fig. 3A). Figure 3B shows the tension force F as a function of the net displacement Δ for a given

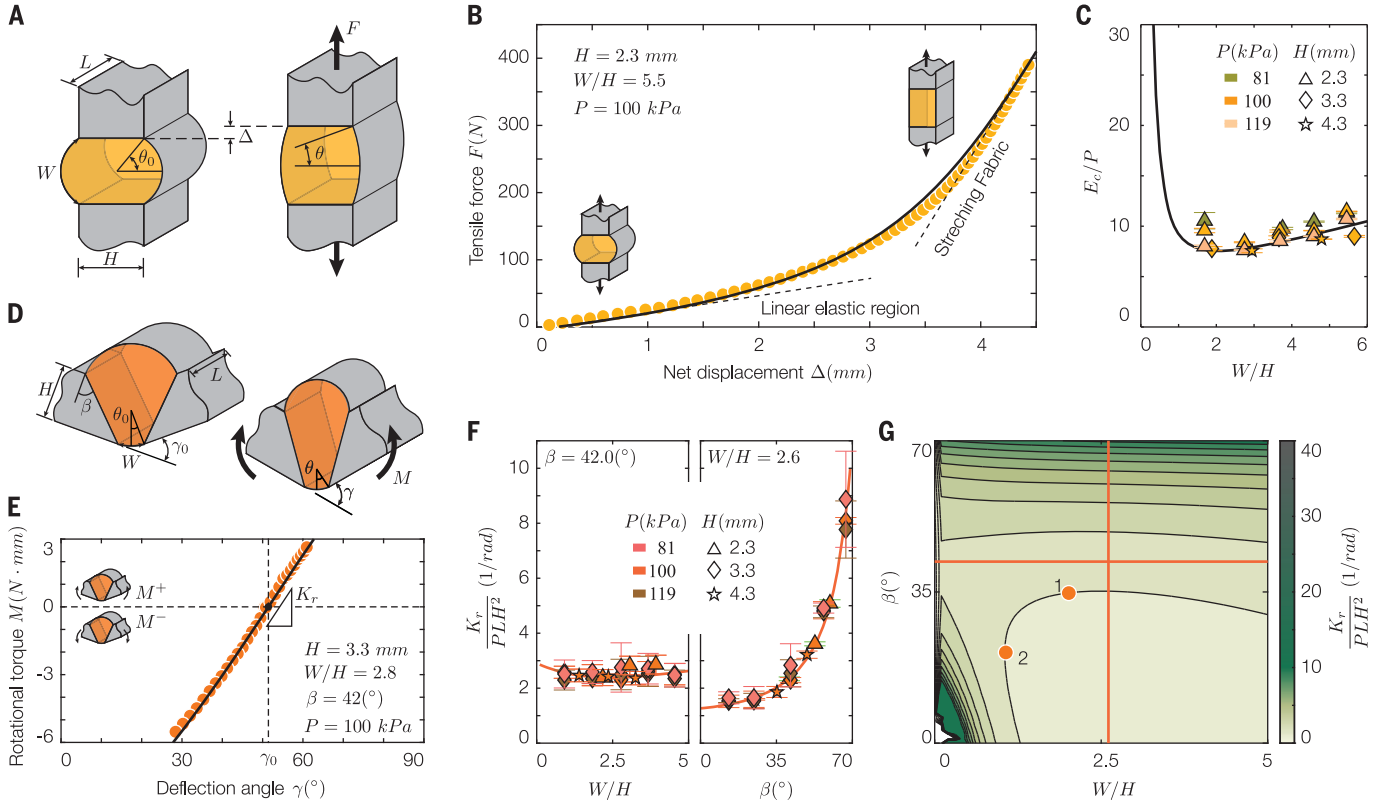


Fig. 3. Mechanical properties of an inflated cell. (**A** to **C**) Stretching modulus. (A) Inflated rectangular cell at rest and under tension with span angle θ and net displacement Δ. (B) Force-displacement curve of an inflated rectangular cell of height H = 2.3 mm and dimensionless width W/H = 5.5 under a pressure P = 10.0 kPa (solid yellow circles). Solid black line indicates theoretical predictions. (C) Dimensionless stretching modulus E_c/P as a function of dimensionless width W/H for various pressures and different cell heights. Solid black line indicates a comparison with the theoretical predictions from Eq. 3. (**D** to **G**) Bending stiffness. (D) Bending test of of an inflated trapezoidal cell. (E) Starting from an equilibrium deflection angle γ_0, applying a torque M results in a new angle γ (H = 3.3 mm, W/H = 2.78, β = 41.5°, P = 10.0 kPa). The slope of the curve in the vicinity of γ_0 = 50.8° corresponds to the rotational stiffness K_r. (F) Rescaled rotational stiffness K_r/PLH^2 as a function of the dimensionless width W/H (β = 42.0°) and of the angle β (W/H = 2.6) at various pressures and different cell heights, respectively. Solid orange lines indicate our nonlinear theoretical prediction (Eq. 4). (G) Contour map of the dimensionless rotational stiffness as a function of W/H and β.

specimen. The theoretical prediction (see the supplementary materials) is plotted in continuous lines and matches precisely the force-displacement curves measured experimentally. For a given applied pressure P, the effective Young modulus E_c of a single cell is inferred from the linear response of the cell for small displacement and is expected to follow

$$E_c = \tan\theta_0 \frac{\cos\theta_0 + \theta_0\sin\theta_0}{\sin\theta_0 - \theta_0\cos\theta_0} P \qquad (3)$$

where the initial spanned angle θ_0 obeys $\theta_0/\cos\theta_0 = W/H$ (see the supplementary materials). The comparison of the normalized modulus E_c/P as a function of the aspect ratio W/H is in excellent agreement with the theoretical prediction, confirming that the modulus only depends on the applied pressure and the geometry of the cell (Fig. 3C). Conversely, for large extension, the stiffness of the cell (Fig. 3B) tends toward the stretching stiffness of the pair of fabric laid flat, $E_{max} = 2E_f t/H$ (where E_f and t are the Young modulus and thickness of the fabric sheet, respectively), which is independent of pressure.

The angular stiffness of a nonsymmetric cell (β ≠ 0°) is obtained by measuring the torque M as a function of the applied deflection angle γ (Fig. 3, D and E). The effective angular stiffness K_r of the cell (linearized response; see the supplementary materials) is predicted to obey

$$\frac{K_r}{PLH^2} = \frac{W}{H}\frac{\cos\left(\theta_0 + \beta - \frac{\gamma_0}{2}\right) + 2\theta_0\sin\left(\theta_0 + \beta - \frac{\gamma_0}{2}\right)}{4\theta_0^2\cos\beta} \qquad (4)$$

where $\theta_0 = \frac{W}{H}\frac{(2\beta-\gamma_0)\cos\beta}{2\sin\beta}$ is the initial value of spanned angle of arc W, L is the length of cell, and the initial folding angle γ_0 follows Eq. 1.

These predictions (continuous line in Fig. 3F) were confirmed by direct experimental measurements of the angular stiffness (with no adjustable parameter). Different configurations (e.g., points 1 and 2 in the contour map in Fig. 3G) are expected to display different actuated geometries but similar structural stiffness for the same pressure. Overall, we found that the mechanical stiffness of an isolated Gaussian cell is directly proportional to its internal pressure, whereas its activated geometry is independent of pressure, suggesting the possibility of tuning the stiffness independently of the shape (for a general discussion of the stiffness of the shell structure as a function of the spatial periodicity of the cells, see the supplementary materials).

Assembling Gaussian cells into programmable structures

We now explore how elementary folding and contraction mechanisms of individual cells can be used to program shape morphing of plates into complex structures. The induced curvature can be discretized as $\kappa = \gamma/\ell$, where ℓ is the spacial periodicity of the cells' distribution (Fig. 2A). Developable surfaces such as generalized cylinders or cones are readily obtained by controlling the deflection angle

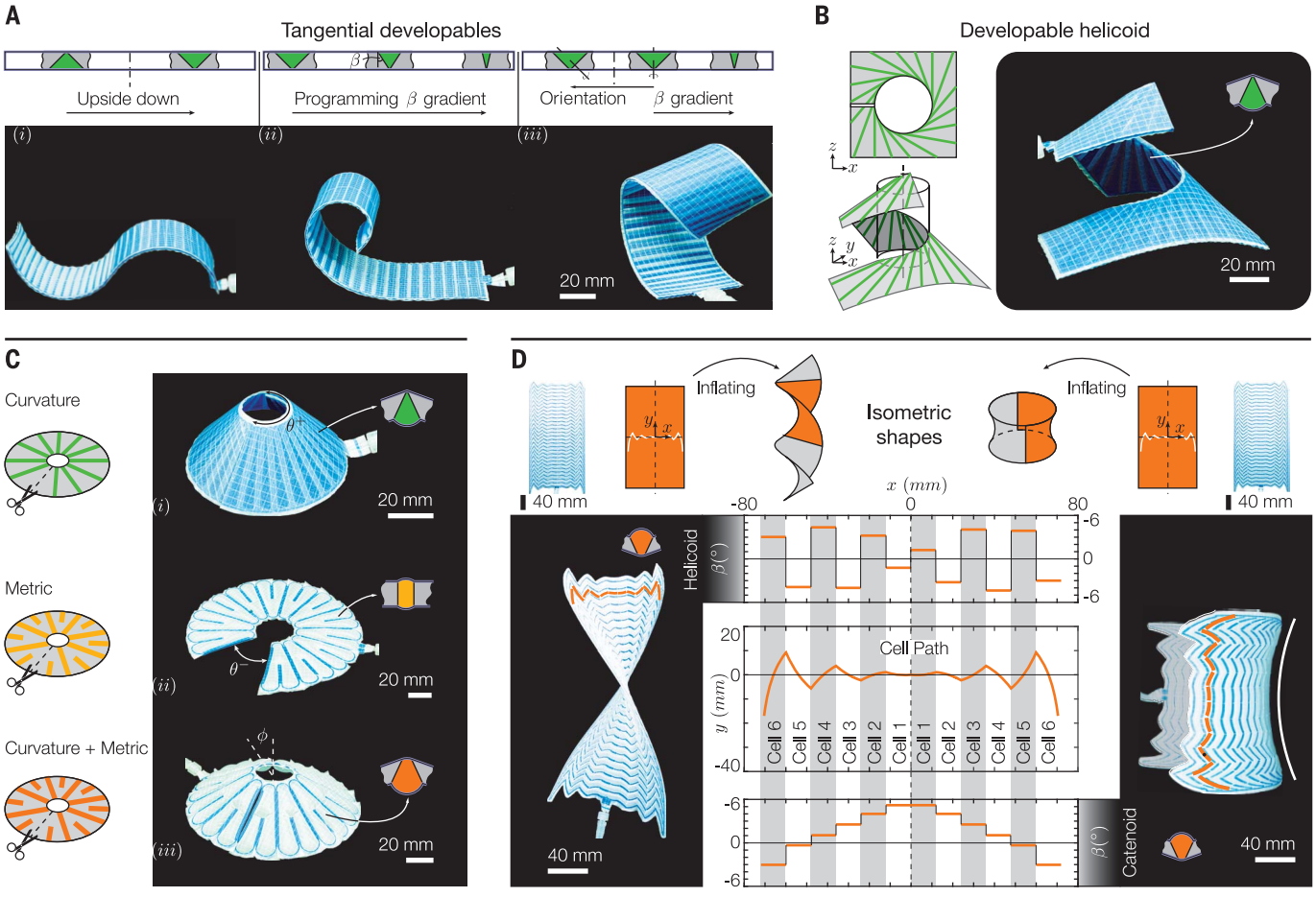


Fig. 4. Programming bending and Gaussian curvatures in multicellular structures. (**A**) Developable shapes based on triangular cells (see movie S2): (i) an "S" shape with opposite curvatures (ii) a spiral with a linearly increasing curvature, and (iii) a curling shape with variable curvature direction. (**B**) Developable helicoid with generating lines tangent to the central circle (see movie S2). (**C**) A sectioned annulus with radial cells can morph into (i) a truncated cone with angular surplus (pure bending with triangular cells), (ii) a flat annulus sector with angular deficit (metric change with rectangular cells), or (iii) a perfect truncated cone when bending and metric changes are programmed simultaneously (see movie S3). (**D**) Selection between isometries (see movie S4). The same metric change (programmed zigzag pattern) leads to degenerated isometric shapes from helicoid to catenoid. Controlling local bending breaks the degeneracy (for a discussion of inverse programming, see the supplementary materials).

locally. Figure 4A illustrates different examples in which the sign of a constant curvature is alternated by piece, the amplitude of the curvature is proportional to the curvilinear abscissa, or the direction of the bending direction is modified, leading, respectively, to (i) an "S" shape, (ii) a spiral with linearly increasing curvature, or (iii) a curly ribbon (see movie S2). A developable helicoid is finally obtained by selecting the director lines tangent to a central circle cut inside the initially flat structure, with an additional transverse cut to allow for deployment (Fig. 4B and movie S2). Figure 4C and movie S3 demonstrate, with the simple example of a slit annulus morphing into a cone, why curvature and metric changes should be used simultaneously. Programming solely curvature without in-plane contraction on such a slit annulus does lead to a cone (Fig. 4Ci), but one with an overlap angle $\theta^+ = 2\pi(1 - \sin\varphi)$, where φ is the cone angle. Conversely, an annulus with cells solely programmed for in-plane contraction will remain flat when actuated with an angular deficit $\theta^- = 2\pi(1 - \lambda_{eff})$, where λ_{eff} is the effective contraction ratio at the scale of the whole structure (Fig. 4Cii). Only when both the active curvature and the in-plane contraction were combined in the programming did we observe that the annulus deployed into a perfect target conical shape (Fig. 4Ciii).

In-plane contraction dictates the Gaussian curvature of the deployed structures. However, targeting this quantity alone only restrains the resulting shape to a family of isometries (*12*). In practice, the observed realization relies on a minimization of the finite bending energy of actual structures, which tends to limit the design space to peculiar solutions among these families (*33*, *34*). Archetypal catenoids and helicoids, minimal surfaces commonly illustrated with soap films, present the same distribution of Gaussian curvature. Merely programming in-plane contraction using cells of rectangular cross-section is therefore not enough to select between configurations. However, adjusting the internal angle of the walls allows for the control of the bending curvature in the direction perpendicular to the cells' path. As a result, for two structures designed with similar cell outlines [the zigzag patterns provide biaxial contraction (*35*)], precise adjustment of the internal angle leads to the desired helicoid (of programmed handiness) or catenoid shapes (*36*) (Fig. 4D and movie S4; for a discussion of the detailed design of the cells, see the supplementary materials). Although Gaussian cells are capable of curvature in the direction

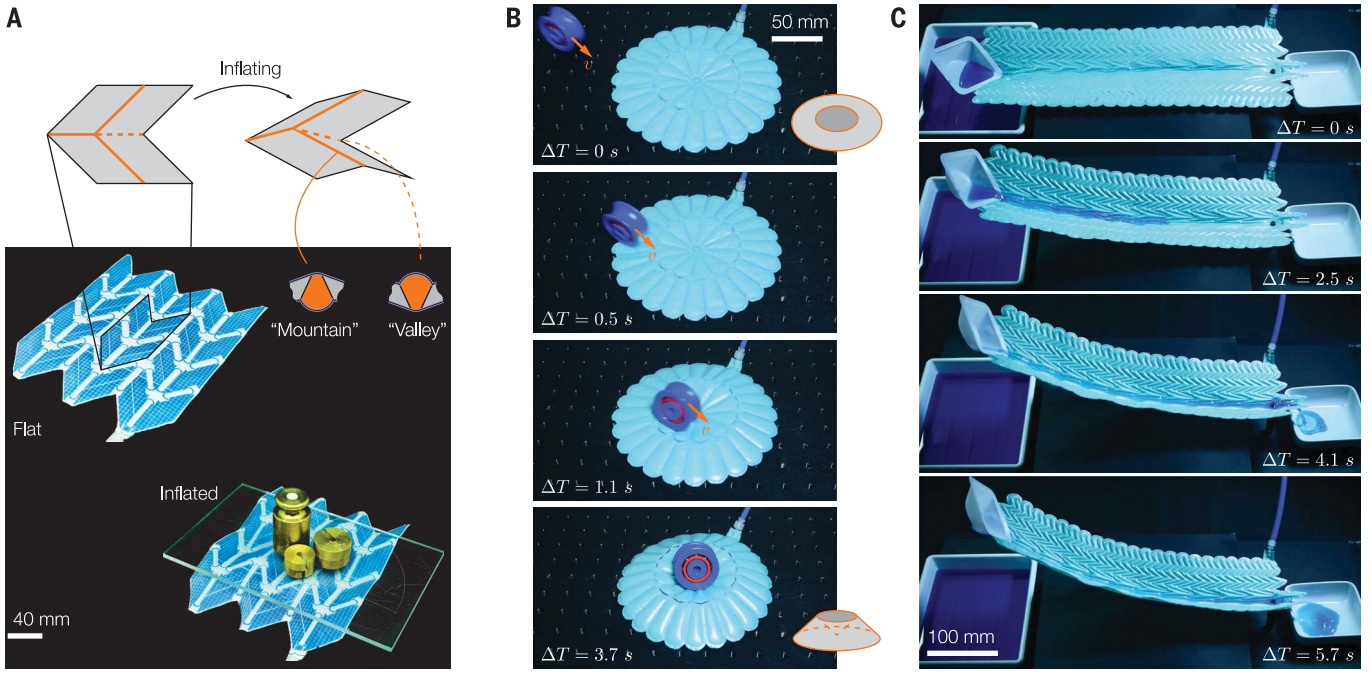


Fig. 5. Toward shape morphing robotics. (**A**) Self-actuated Miura-ori panel (see movie S5). Cells function as stiff "mountain" or "valley" creases (total load of 1 kg). (**B**) Volcano-shaped actuator (see movie S6). The initially flat disk is actuated in a fraction of second as a cylinder rolls over it at 10 cm/s. The cylinder (weight 40 g) remains trapped inside the deployed crater. (**C**) Upon inflation, a flat stripe bends into a doubly curved gutter track that permits manipulation and transport of liquids (see movie S7).

perpendicular to their axis, the use of alternating directions overcomes this constraint.

Much effort has been devoted recently to the design of origami structures (*25*, *26*); however, robust self-actuation remains an issue (*37*, *38*). Figure 5A presents a classical Miura-ori tessellation pattern (*39*) in which "mountain" and "valley" creases have been programmed using Gaussian cells. Upon inflation, the structure self-folds to the target configuration with a stiffness high enough to support a relatively heavy load (see movie S5). Beyond origami structures, their excellent stiffness, robustness to failure, and fast actuation suggest that pneumatic Gaussian cells are promising for applications in morphing robotics (for a discussion of performances and scalability, see the supplementary materials). Standard soft-robotics technologies generally rely on bending actuation of grabbing beam elements (*11*); however, robotic systems could be developed by harnessing the surface shape morphing achieved with Gaussian cells. Figure 5B illustrates an initially flat disk that adopts in a fraction of second the shape of a crater based on the cone transformation and traps an object rolling over (see movie S6, which also demonstrates how a ball can be guided by an active track). Although the presented examples are based on a single inlet and interconnected cells, more complex kinematics could be obtained by separating cells into subgroups that would be activated separately or by embedding mechanical valves between cells that could provide a sequential actuation (*40*, *41*). Finally, active surfaces may also be used to manipulate liquids. In Fig. 5C and movie S7, a long strip is programmed to morph into a curved gutter of positive Gaussian curvature (*35*) with a controlled bending direction, transferring liquid to a different container and demonstrating the versatile capabilities of the technique.

Toward shape-morphing robotics

Inspired by actuating cells observed in some plants, we have developed powerful and versatile Gaussian cells that provide stiff, fast, and reversible self-morphing of thin plates. Because metrics and curvature can be programmed simultaneously, these cells provide accurate control of the deployed geometry, imparting a high degree of mechanical stability to the resulting structure. Moreover, the stiffness of the deployed shell (proportional to applied pressure) can be tuned independently of its shape (dictated by the local configuration), opening up the possibility of developing large-scale morphing structures. Pneumatic Gaussian cells combine stiffness with a high level of control in shape programming, holding significant promise for applications in which avoiding mechanical instabilities is crucial, but also for haptic devices in which both shape and stiffness controls are valuable output information for the end user. Made using standard 3D printing techniques with consumer-grade materials, the resulting robust, shape-changing surfaces bring the concept of material-machine closer to providing a large-scale and general commodity product.

REFERENCES AND NOTES

1. H. Liang, L. Mahadevan, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 5516–5521 (2011).
2. X. Cheng, Y. Zhang, *Adv. Mater.* **31**, e1901895 (2019).
3. L. Hines, K. Petersen, G. Z. Lum, M. Sitti, *Adv. Mater.* **29**, 1603483 (2017).
4. T. J. Jones, E. Jambon-Puillet, J. Marthelot, P.-T. Brun, *Nature* **599**, 229–233 (2021).
5. R. F. Shepherd *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 20400–20403 (2011).
6. J. Wang, D. Gao, P. S. Lee, *Adv. Mater.* **33**, e2003088 (2021).
7. M. Li, A. Pal, A. Aghakhani, A. Pena-Francesch, M. Sitti, *Nat. Rev. Mater.* **7**, 235–249 (2022).
8. Y. Yang, K. Vella, D. P. Holmes, *Sci. Robot.* **6**, eabd6426 (2021).
9. W. Kim *et al.*, *Sci. Robot.* **4**, eaay3493 (2019).
10. D. Melancon, B. Gorissen, C. J. García-Mora, C. Hoberman, K. Bertoldi, *Nature* **592**, 545–550 (2021).
11. B. Gorissen *et al.*, *Adv. Mater.* **29**, 1604977 (2017).
12. D. J. Struik, *Lectures on Classical Differential Geometry* (Addison-Wesley Press, 1950).
13. J. Dervaux, M. Ben Amar, *Phys. Rev. Lett.* **101**, 068101 (2008).
14. A. B. Rebocho, P. Southam, J. R. Kennaway, J. A. Bangham, E. Coen, *eLife* **6**, e20156 (2017).
15. Y. Klein, E. Efrati, E. Sharon, *Science* **315**, 1116–1120 (2007).
16. E. Efrati, E. Sharon, R. Kupferman, *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* **80**, 016602 (2009).
17. J. Kim, J. A. Hanna, M. Byun, C. D. Santangelo, R. C. Hayward, *Science* **335**, 1201–1205 (2012).
18. A. Nojoomi, J. Jeon, K. Yum, *Nat. Commun.* **12**, 603 (2021).
19. H. Aharoni, Y. Xia, X. Zhang, R. D. Kamien, S. Yang, *Proc. Natl. Acad. Sci. U.S.A.* **115**, 7206–7211 (2018).
20. T. van Manen, S. Janbaz, A. A. Zadpoor, *Mater. Horiz.* **4**, 1064–1069 (2017).
21. J. Gu *et al.*, "Geodesy: Self-rising 2.5D tiles by printing along 2D geodesic closed path," in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (ACM, 2019), pp. 1–10.

22. J. H. Pikul *et al*., *Science* **358**, 210–214 (2017).
23. E. Siéfert, E. Reyssat, J. Bico, B. Roman, *Nat. Mater.* **18**, 24–28 (2019).
24. E. Siéfert, E. Reyssat, J. Bico, B. Roman, *Soft Matter* **16**, 7898–7903 (2020).
25. L. H. Dudte, E. Vouga, T. Tachi, L. Mahadevan, *Nat. Mater.* **15**, 583–588 (2016).
26. S. J. Callens, A. A. Zadpoor, *Mater. Today* **21**, 241–264 (2018).
27. G. P. T. Choi, L. H. Dudte, L. Mahadevan, *Nat. Mater.* **18**, 999–1004 (2019).
28. A. S. Gladman, E. A. Matsumoto, R. G. Nuzzo, L. Mahadevan, J. A. Lewis, *Nat. Mater.* **15**, 413–418 (2016).
29. J. W. Boley *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **116**, 20856–20862 (2019).
30. S.-J. Jeon, R. C. Hayward, *Soft Matter* **16**, 688–694 (2020).
31. S. Matschi *et al*., *Plant Direct* **4**, e00282 (2020).
32. A. Mader, M. Langer, J. Knippers, O. Speck, *J. R. Soc. Interface* **17**, 20200358 (2020).
33. E. Efrati, E. Sharon, R. Kupferman, *J. Mech. Phys. Solids* **57**, 762–775 (2009).
34. S. Armon, E. Efrati, R. Kupferman, E. Sharon, *Science* **333**, 1726–1730 (2011).
35. T. Gao, E. Siéfert, A. DeSimone, B. Roman, *Adv. Mater.* **32**, e2004515 (2020).
36. H. Aharoni, E. Sharon, R. Kupferman, *Phys. Rev. Lett.* **113**, 257801 (2014).
37. C. D. Santangelo, *Annu. Rev. Condens. Matter Phys.* **8**, 165–183 (2017).
38. S. Wu, T. Chen, H. Tsai, *J. Mech.* **35**, 627–639 (2019).
39. M. Schenk, S. D. Guest, *Proc. Natl. Acad. Sci. U.S.A.* **110**, 3276–3281 (2013).
40. L. Jin, A. E. Forte, K. Bertoldi, *Adv. Sci. (Weinh.)* **8**, e2101941 (2021).
41. L. C. van Laake, J. de Vries, S. M. Kani, J. T. Overvelde, *Matter* **5**, 2898–2917 (2022).
42. S. D. Guest, S. Pellegrino, "Inextensional wrapping of flat membranes," in *Proceedings of the First International Seminar on Structural Morphology*, R. Motro, T. Wester, Eds. (University of Montpellier, 1992), vol. 25, pp. 203–215.
43. D. Webb *et al*., "Starshade mechanical architecture technology effort," paper presented at the 3rd AIAA Spacecraft Structures Conference, San Diego, CA, 4–8 January 2016.

ACKNOWLEDGMENTS

We thank H. Aharoni, E. Reyssat, and A. DeSimone for stimulating discussions and B. Moulia for botanical inspiration. **Funding:** This work was partly supported by the French ANR SecondSkin project (grant ANR-21-CE33-0018). **Author contributions:** T.G., J.B., and B.R. conceived and developed the pneumatic Gaussian cell concept. T.G., J.B., and B.R. conducted the experiments and analyzed the data. T.G., J.B., and B.R. developed the theoretical model. All authors wrote the manuscript. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data are available in the main text or supplementary materials. **License information:** Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adi2997
Materials and Methods
Supplementary Text
Figs. S1 to S20
References
Movies S1 to S7

Submitted 17 April 2023; accepted 20 July 2023
10.1126/science.adi2997

SOLAR PHYSICS

Picoflare jets power the solar wind emerging from a coronal hole on the Sun

L. P. Chitta[1]*, A. N. Zhukov[2,3], D. Berghmans[2], H. Peter[1], S. Parenti[4], S. Mandal[1], R. Aznar Cuadrado[1], U. Schühle[1], L. Teriaca[1], F. Auchère[4], K. Barczynski[5,6], É. Buchlin[4], L. Harra[5,6], E. Kraaikamp[2], D. M. Long[7,8], L. Rodriguez[2], C. Schwanitz[5,6], P. J. Smith[7], C. Verbeeck[2], D. B. Seaton[9]

Coronal holes are areas on the Sun with open magnetic field lines. They are a source region of the solar wind, but how the wind emerges from coronal holes is not known. We observed a coronal hole using the Extreme Ultraviolet Imager on the Solar Orbiter spacecraft. We identified jets on scales of a few hundred kilometers, which last 20 to 100 seconds and reach speeds of ~100 kilometers per second. The jets are powered by magnetic reconnection and have kinetic energy in the picoflare range. They are intermittent but widespread within the observed coronal hole. We suggest that such picoflare jets could produce enough high-temperature plasma to sustain the solar wind and that the wind emerges from coronal holes as a highly intermittent outflow at small scales.

Hot plasma continuously escapes the Sun into the heliosphere, forming the solar wind (*1*, *2*). Fast solar wind streams (those with speeds >500 km s^{-1}) have been traced to large (mostly polar) coronal holes—features formed by open magnetic fields on the Sun that appear dark in coronal images taken at extreme ultraviolet (EUV) or x-ray wavelengths (*3*–*5*). Small coronal holes that form adjacent to active regions (areas on the Sun with strong magnetic fields, including sunspots) could be the source regions of slower wind, with speeds <500 km s^{-1} (*6*, *7*).

The physical origin and acceleration mechanism of solar wind from coronal holes are debated (*5*, *8*, *9*). Potential drivers include wave dissipation and turbulent cascades (*10*–*12*) or interchange reconnection between open and closed magnetic field lines at the coronal base (*13*–*15*). Jets with temperatures of ~0.1 MK emerge from the transition region network at the base of the corona (network jets), which are common in coronal hole regions, and could sustain the solar wind (*16*). Coronal counterparts to network jets are rare (*17*, *18*), so it is unclear whether the cooler network jets can efficiently supply hot plasma to drive the solar wind.

Previous EUV imaging has shown signatures of jet outflows from plumes (bright ray-like features embedded in coronal holes) and the ambient, interplume coronal hole regions (*19*, *20*). EUV images, with spatial resolution of ~1000 km, have shown that plumes have substructures consisting of jets with widths as small as 1500 km (*21*–*23*). These jets appear to be universal in the corona: They emerge from coronal hole plumes, from quiet regions and from active regions of the Sun's atmosphere (*24*). Jets from plumes could sustain the solar wind mass flux during the minimum of the 11-year solar activity cycle and up to 60% of the mass flux during times of maximum activity (*23*). Because plumes fill only 10% of a coronal hole's volume (*25*), the bulk of the wind might originate in interplume regions (*26*–*28*). However, the spatial structure and timescales of interplume outflows are not known.

The solar wind emerging from coronal holes connects the Sun to the heliosphere. Spacecraft measurements in the inner heliosphere have been interpreted as showing that reversals of the radial magnetic field (known as switchbacks) in the solar wind are modulated on spatial scales as small as 1 Mm (the typical convective granular scale of the photosphere) at the base of the Sun's corona (*29*). However, EUV imaging has not been available with subgranular resolution (<1 Mm), which is necessary to investigate this proposal.

EUV observations of a coronal hole

We observed a coronal hole using the EUV high-resolution imager (HRI_{EUV}), which is part of the Extreme Ultraviolet Imager (EUI) (*30*) instrument on the Solar Orbiter spacecraft (*31*). The passband of HRI_{EUV} is centered at 17.4 nm, with its thermal response function peaking at temperatures of ~1 MK because of emission lines of Fe IX (at 17.11 nm) and Fe X (at 17.45 and 17.72 nm).

The coronal hole was located close to the Sun's south pole. The data were obtained on 30 March 2022 between 04:30 and 05:00 universal time (UT) as part of an observing campaign performed around perihelion (the closest approach of the spacecraft to the Sun). The images have a spatial resolution of ~237 km on the Sun and a cadence of 3 s (*32*).

We split the coronal hole into three regions (labeled R1 to R3 in Fig. 1A) to investigate the small-scale coronal dynamics. Regions R1 and R3 are darker portions of the coronal hole with a few embedded brighter emission structures, known as coronal bright points, whereas R2 is an isolated plume within the coronal hole.

Y-shaped transient jets

We identify a Y-shaped jet with a base extent of ~5 Mm (between boxes R1-2 and R1-4 in Fig. 1B). Jets with such Y-shaped morphology are common in coronal holes (*33*). We focus on smaller jets, on spatial scales of <1 Mm, emerging from the ambient coronal hole. We visually identified 10 such small-scale jets within R1 (Fig. 1, C to L). Some of these jets show a full or partial Y-shaped morphology (jets 1, 2, 3, 6, 7, 8, and 10), whereas others show a much simpler linear morphology (jets 4, 5, and 9). It is possible that all are intrinsically Y shaped, but the Y-shaped structure is not visible in some cases because of the viewing angle, which could lead to both legs of the jet being aligned with the line of sight. All of these jets have fine-scale structure down to the spatial resolution of our images.

These small jets exhibit intensity enhancements, over the background emission from the coronal hole. Figure 2 shows light curves of three example jets in region R1. For each of these, we selected two spatial points along the path of the jet to study the intensity variations (fig. S1). The intensity enhancements (defined as the amplitude of intensity fluctuations compared with the local background) are in the range of 100 to 300 digital numbers (DN) per pixel (equivalent to counts of ~15 to 44 photons per pixel; for comparison, coronal bright points in our observations are ~500 photons per pixel). Depending on the number of pixels used for spatial averaging of the signal, peak intensities of all these jets (above the background average intensity) are at least three times the local background photon noise (*32*).

The observed intensity fluctuations vary on timescales of 20 to 60 s (Fig. 2). Features associated with an intense jet base (Fig. 1, E to H) at least partially persist for longer timescales of 300 to 600 s but also exhibit intensity fluctuations on shorter timescales of 100 s. We estimate that the jet speeds are ~100 km s^{-1}, with the speeds of different jets varying by a factor of 2 (Fig. 2, D to F) (*32*). These differences in measured jet speeds could be a result of projection effects. We observed repeated jet activity in these regions and visually counted a total of 38 small jets from the 10 locations

[1]Max-Planck-Institut für Sonnensystemforschung, 37077 Göttingen, Germany. [2]Solar-Terrestrial Centre of Excellence, Solar Influences Data Analysis Centre, Royal Observatory of Belgium, 1180 Brussels, Belgium. [3]Skobeltsyn Institute of Nuclear Physics, Moscow State University, Moscow 119991, Russia. [4]Institut d'Astrophysique Spatiale, Centre National de la Recherche Scientifique, Université Paris-Saclay, 91405 Orsay, France. [5]Physikalisch-Meteorologisches Observatorium Davos, World Radiation Center, 7260 Davos Dorf, Switzerland. [6]Eidgenössische Technische Hochschule Zürich, 8093 Zürich, Switzerland. [7]Mullard Space Science Laboratory, University College London, Dorking, Surrey RH5 6NT, UK. [8]Astrophysics Research Centre, School of Mathematics and Physics, Queen's University Belfast, Belfast BT7 1NN, Northern Ireland, UK. [9]Southwest Research Institute, Boulder, CO 80302, USA.
*Corresponding author. Email: chitta@mps.mpg.de

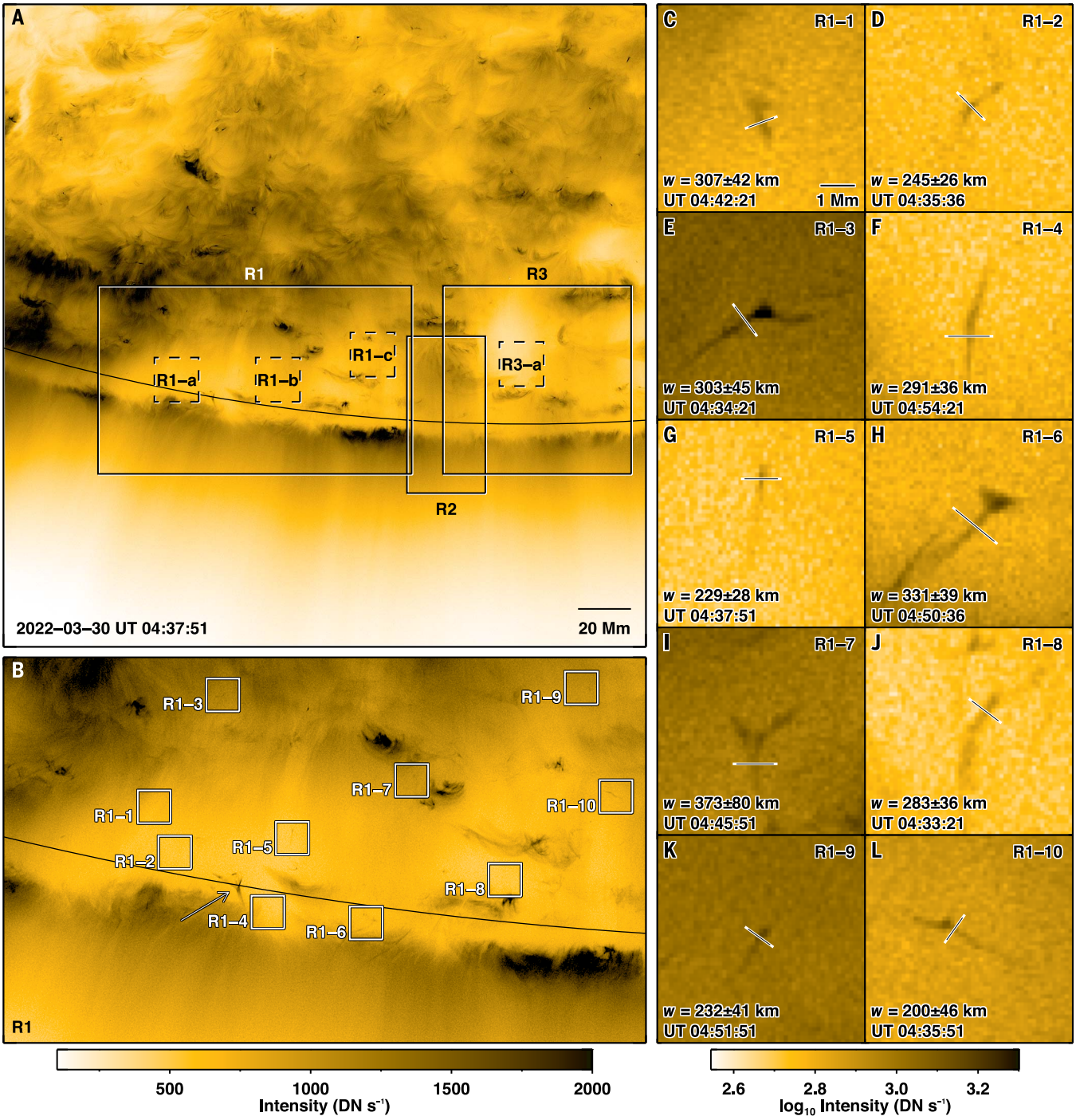


Fig. 1. EUV images of jets from a coronal hole. All images are displayed on a negative color scale (i.e., black is bright). Each panel shows data with a 15-s effective cadence (*32*), and south is roughly downward. (**A**) An overview HRI_{EUV} image of a south polar coronal hole recorded on 30 March 2022 around 04:38 UT. The field of view is ~243 Mm × 243 Mm. Solid boxes labeled R1 to R3 outline regions of the coronal hole discussed in the text. Dashed boxes labeled R1-a, R1-b, R1-c, and R3-a are interplume regions. The solid curve indicates the position of the Sun's limb. (**B**) Zoomed-in view of R1 from (A) on the same brightness scale. Boxes labeled R1-1 to R1-10, each ~6 Mm × 6 Mm, indicate locations within the coronal hole where we identified small-scale jet activity. The arrow points to an ~5-Mm-scale–larger Y-shaped jet. An animated version of this panel is shown in movie S1. (**C** to **L**) Closer views of the boxes in (B) on a logarithmic brightness scale. White lines indicate locations where the width of each jet *w* was measured; its value is shown at the bottom of each panel (also listed in table S1).

labeled in Fig. 1 over the course of 30 min of observations. We also identified 29 additional small-scale jets in the wider coronal hole region (movie S1).

Plume-related jets

The emission from region R2 (Fig. 1A) is enhanced over the immediate surroundings, which suggests that it is a coronal hole plume. We observed clusters of small jets emerging throughout the associated plume region (Fig. 3). These plume jets do not have the Y-shaped morphology of the jets discussed above.

Previous lower-resolution EUV observations have shown that outflows from plumes are structured in the form of jets (*21*) and that the outflows are quasiperiodic in nature, with speeds of ~100 km s^{-1} (*19*, *23*). Our higher-resolution HRI_{EUV} observations show even smaller jets, with cross-sectional widths close to the spatial resolution limit of ~200 km (a very thin jet is shown in Fig. 3H). We observe at least 30 such jets from the plume (movie S2).

Faint interplume emission and jets

Although both the interplume Y-shaped jets and the plume jets are spatially discrete, we found that there is a more widespread emission that is faint yet above the background level. This fainter emission, with veil-like morphology, exhibits apparent outward propagation throughout the coronal hole (fig. S3 and movies S3 and S4). The fainter emission also originates from the interplume regions; we designate four examples as R1-a, R1-b, R1-c, and R3-a (Fig. 1A). These regions of fainter emission are partially obscured by the foreground emission in R1, but the source region in R3 is unobscured.

This fainter, veil-like emission is also spatially structured. There are outflows apparently emerging from the identified source regions in the form of small-scale jets. Figure 4 shows three example jets from three of the identified interplume sources. Compared with the Y-shaped jets in Fig. 1, these interplume jets are even fainter; however, they are still at least three times the local background photon noise above the background mean intensities. These fainter jets exhibit intensity enhancements of ~50 DN per pixel over the local background (Fig. 4 and fig. S2). These interplume jets have similar morphology to that of the plume counterparts, and their cross-sectional widths are close to the resolution limit of the images. Interplume jets also appear in clusters on larger scales—>10 Mm—similar to plume jets. They are associated with intensity propagations; although we cannot determine these jet speeds because of noise, we constrain them to be supersonic outflows >100 km s^{-1} (*32*).

Magnetic origins of the jets

The high speeds, short timescales, and narrow widths of the observed jets indicate that they are driven by a variable energy source located at the coronal base, on spatial scales of $\lesssim$ 100 km. Y-shaped jets (Fig. 1) are already known to be driven by reconnection of open and closed magnetic field lines (*34*–*36*). The base width of $\lesssim$ 1 Mm indicates that the observed small, Y-shaped jets are associated with the emergence or cancellation of granular-scale magnetic features (*37*). Although the plume and fainter interplume jets we observed on the solar disc do not have Y-shaped morphology, similar features observed on the Sun's limb do have morphologies consistent with the magnetic topology of interchange reconnection between open and closed magnetic domains (figs. S4 to S6 and movies S5 to S9). We therefore suggest that all the jets we observed are transiently driven by reconnection.

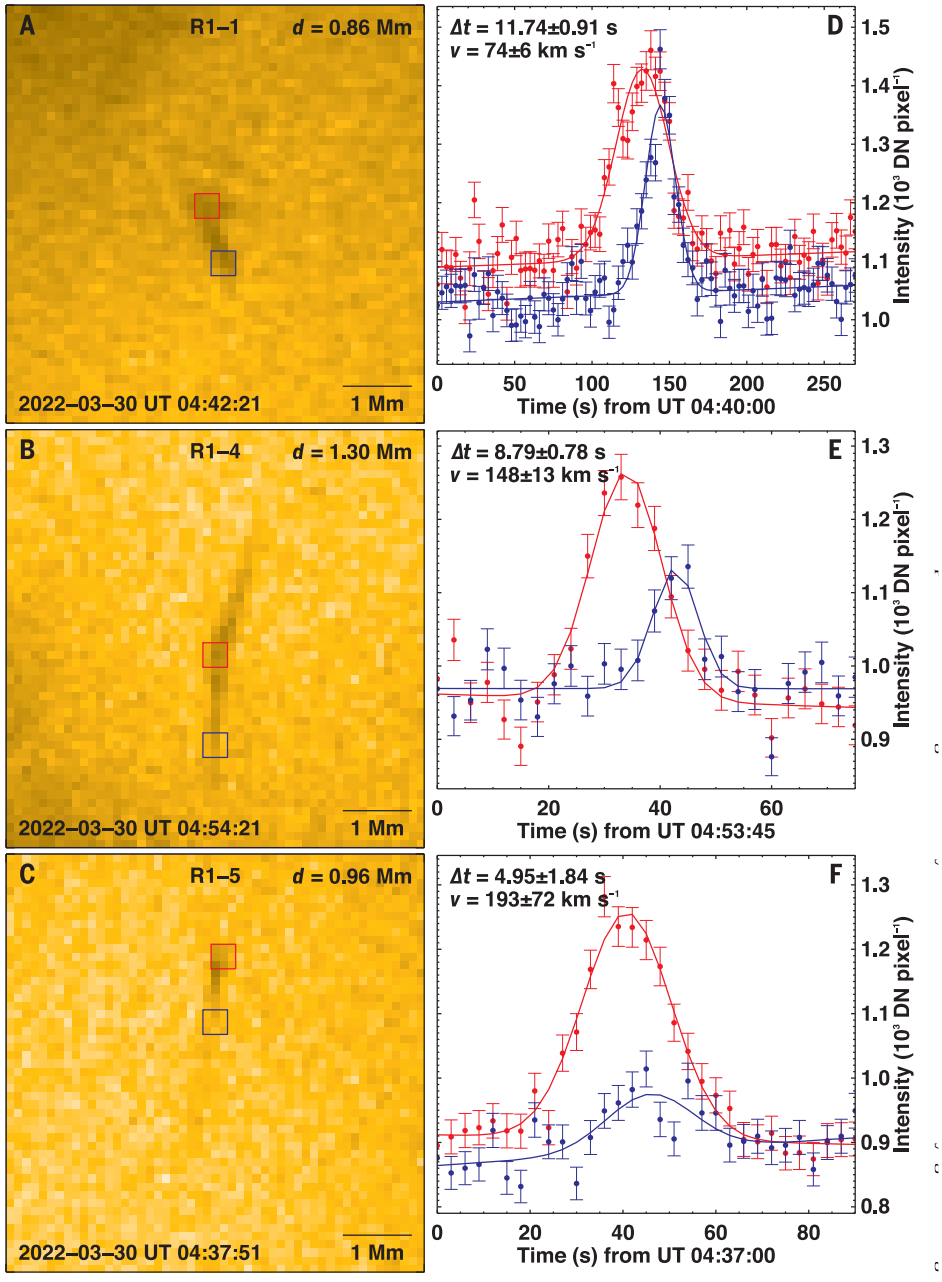


Fig. 2. Temporal variability in three example small-scale jets. (**A** to **C**) Same EUV images as Fig. 1, C, F, and G, respectively. The red and blue squares mark locations along each jet, separated by a plane-of-sky distance *d* listed in the top right of each panel. (**D** to **F**) Data points are average intensities within the corresponding red and blue boxes in (A) to (C) as a function of time, with a cadence of 3 s. Error bars indicate 1σ uncertainties because of photon Poisson (shot) noise (*32*). The red and blue solid curves are Gaussian models fitted to the observed intensities. The separation between the centroids of the two Gaussian peaks Δt and the speed $v = d/\Delta t$ (with their standard errors) are listed in each panel (also listed in table S2).

Previous observations have shown that the Ne VIII emission line, which samples transition region and coronal plasma with temperatures of ~0.63 MK and above, has systematic blueshifts indicating upward flows of ~10 km s^{-1} in coronal holes (*13*, *38*). The outflow speeds in interplume regions can be up to 70 km s^{-1} (*27*, *28*). The small-scale jets that we observed could be the ~1-MK counterparts of the coronal hole outflows previously inferred through spectroscopy at lower spatial resolution. The high-speed jets driven by reconnection in our

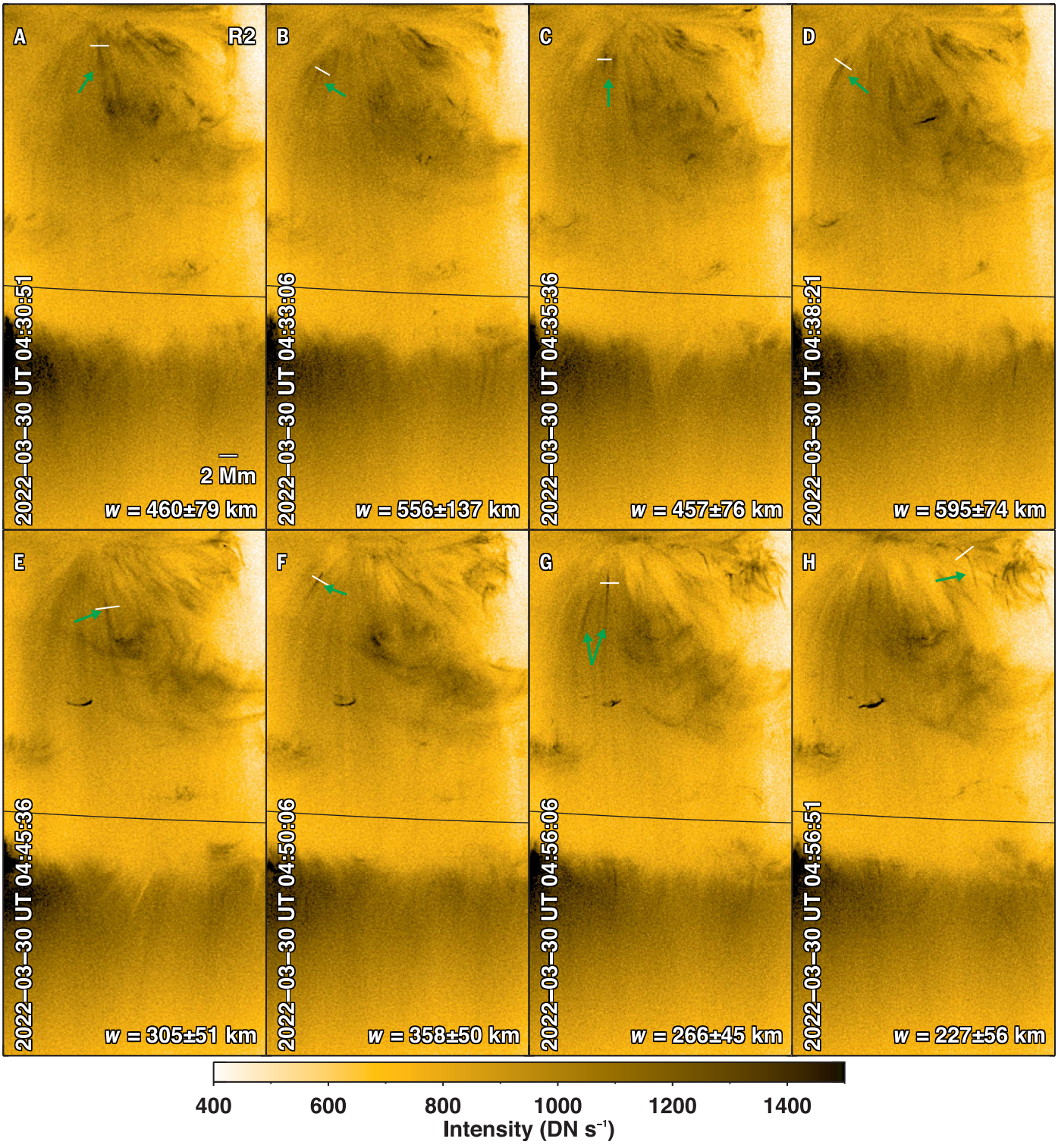


Fig. 3. Jets from plumes. (**A** to **H**) The panels all show region R2 from Fig. 1A but for different times (labeled on each image) with an effective cadence of 15 s (*32*). Small-scale jet activity is indicated by green arrows. White lines indicate locations where the cross-sectional width *w* of each jet was measured, with the results listed in each panel (also listed in table S3). The solid curve marks the solar limb. An animated version of this figure is shown in movie S2.

observations propagate outward and could channel some, if not all, of the material contained within the jets along the open magnetic field lines of the coronal hole, powering the solar wind.

We do not observe any signatures of downward flows in the jets. If they exist, they could be too weak to be detectable in our observations, or they might contain plasma at different temperatures to which the HRI_{EUV} instrument is not sensitive.

Jet energetics

Assuming that the apparent outflows we detected are mass motions, we estimate a lower limit on the kinetic energy flux of $\gtrsim 10^5$ erg cm^{-2} s^{-1} (*32*). This is probably an

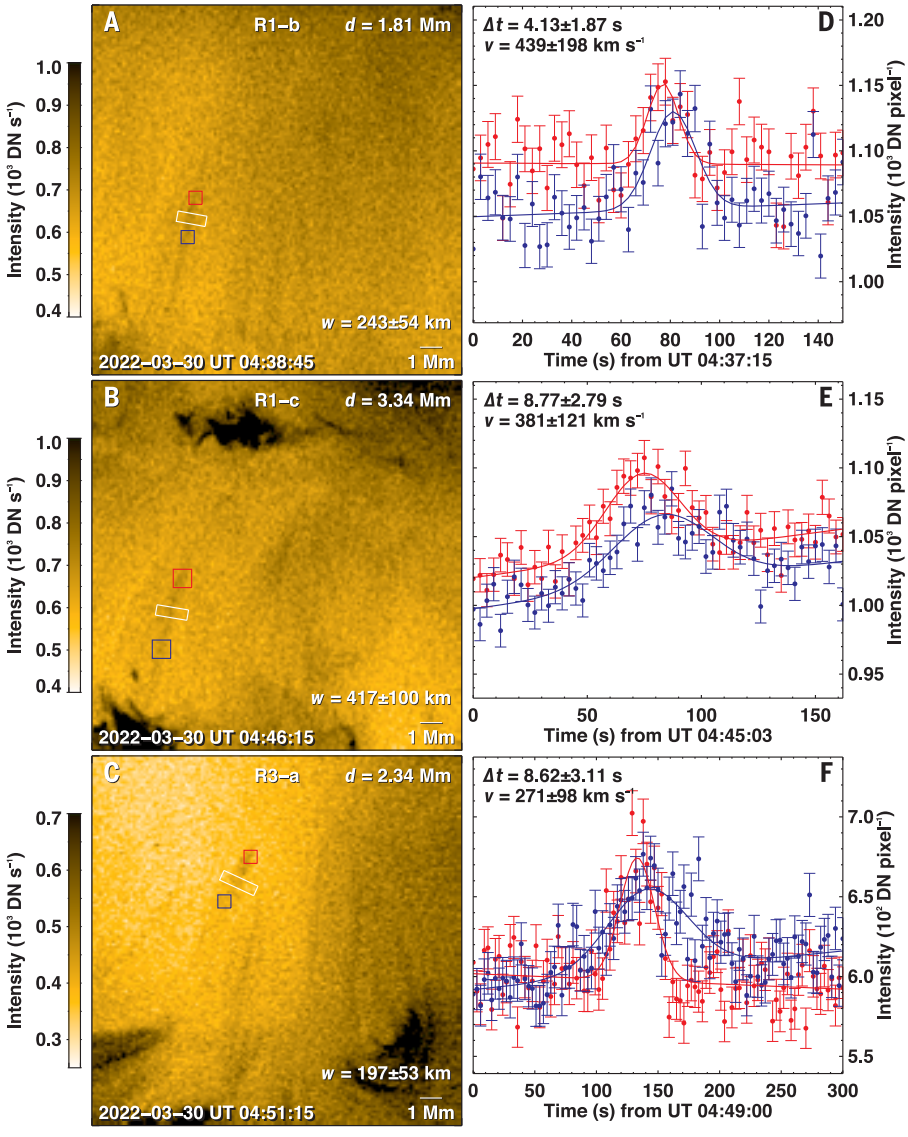


Fig. 4. High-speed faint jets from interplume regions. (**A** to **F**) Same as Fig. 2 but for the boxes labeled R1-b, R1-c, and R3-a in Fig. 1A. Cross-sectional width *w* of each jet was measured by averaging the intensity across the white strip, with the results listed in each panel. The images in (A) to (C) have an effective cadence of 30 s. The average intensities in (D) to (F) have a cadence of 3 s (*32*). Jet parameters are listed in table S4.

underestimate because of line-of-sight effects. Increasing the jet speed by a factor of 2 to 4 would increase the kinetic energy flux of a given jet by nearly one to two orders of magnitude.

Our estimate of the kinetic energy flux in these jets is similar to the minimum nonradiative energy input required to power magnetically open coronal regions (*39*). The narrowest jets we observed have cross-sectional widths of only a few pixels, corresponding to diameters of <237 km, assuming a cylindrical shape. The observed intensity fluctuations are on timescales of 20 to 100 s. If we take that to be the lifetime of the jet and combine it with the derived kinetic energy flux, we infer that the lower limit for the kinetic energy content of a single narrow, faint jet is $\gtrsim 10^{21}$ erg. The typical energy content of a nanoflare is 10^{24} erg (*40*), so the kinetic energy content of fainter interplume jets is in the picoflare range. We therefore refer to these features as picoflare jets. The Y-shaped jets (Fig. 1) and the plume jets (Fig. 3), which are somewhat brighter, presumably carry more kinetic energy.

Plasma outflows from picoflare jets, channeled along the open magnetic field lines of coronal holes, could contribute to the mass flux of the solar wind. We estimate the total mass-loss rate $\dot{m}$ of the observed jets as $\dot{m} = 4\pi R_{\odot}^2 \rho v f_{CH} f_j$, where $R_{\odot} = 695$ Mm is the radius of the Sun, $\rho = 3.34 \times 10^{-16}$ g cm^{-3} is the jet mass density, $v = 100$ km s^{-1} is the jet speed (*32*), $f_{CH} \approx 0.1$ is the fractional area coverage of coronal holes on the Sun (*41*), and f_j is the filling factor of picoflare jets. During periods of high solar activity, upward flows from the edges of active regions could also contribute to the solar wind mass flux (*42*–*45*). If picoflare jets account for the wind mass flux of ~10^{12} g s^{-1} throughout the solar cycle (*46*), their filling factor f_j would need to be ~0.05.

Our observations do not spatially resolve the smallest jets (*32*), especially those in the interplume regions that have weaker intensity enhancements and lower contrast with respect to the local background. Our results are consistent with previous EUV imaging at lower spatial resolution (*24*). The picoflare jets are common in our images (figs. S7 and S8), particularly in the interplume regions that cover a substantial portion of the coronal hole (movies S10 to S12). Above the limb, we observe persistent jet activity that fills the plane of sky (fig. S9). Any fainter jets would not be distinguishable from the background intensity fluctuations in our data.

Implications for the solar wind

Given their frequency in our observations, we suggest that these picoflare jets might be ubiquitous. If so, those emerging from the interplume regions, and to a lesser extent those from plume regions, could provide substantial plasma and energy flux to the solar wind throughout the solar cycle. This is consistent with previous suggestions that small jets make a substantial contribution to the solar wind (*24*).

The intermittent variability and short lifetimes (20 to 100 s) of the observed jets are consistent with the inherently ephemeral nature of magnetic reconnection. Our observations are not sensitive to any steady component of the solar wind. Nevertheless, we suggest that a dynamic component of the solar wind emerges from coronal holes as an intermittent outflow at the coronal base, which is observed as picoflare jets.

Neighboring small-scale, high-speed outflows with different speeds would produce a velocity shear, leading to instabilities (*47*, *48*). We suggest that this could occur in both plume and interplume regions (*32*). If picoflare jets are ubiquitous, these instabilities could play a role in the formation of structures in the solar wind, such as magnetic switchbacks (*7*). In situ spacecraft measurements of the solar wind in the inner heliosphere have been interpreted as showing that the sources responsible for the large- and midscale modulation of switchbacks operate on spatial scales similar to the typical convective scales of granules and supergranules on the Sun (*29*). The narrow widths of the picoflare jets that we observed are compatible with them having underlying source regions of <1 Mm, which is smaller than the typical granule scale (*49*). The larger clusters of picoflare jets that we observed have structure

on scales >10 Mm, which is close to the supergranule scales of 10 to 60 Mm (*50*). Our interpretation that the picoflare jets are driven by granular-scale reconnection could therefore be related to the mid- to large-scale modulations of switchbacks in the solar wind.

The lifetimes and speeds of the picoflare jets are similar to those of network jets, which are common in the transition region (*16*), so the phenomena might be related. Our data cannot determine whether a hot picoflare jet (~1 MK) hosts plasma that has been heated from a cooler network jet phase or whether there are two populations of jets present simultaneously (colder network jets and hotter coronal picoflare jets). We suggest that picoflare jets could supply heated material to the solar wind as could the transition region jets if their energy is dissipated (*16*).

REFERENCES AND NOTES

1. E. N. Parker, *Astrophys. J.* **128**, 664 (1958).
2. M. Neugebauer, C. W. Snyder, *Science* **138**, 1095–1097 (1962).
3. J. Woch *et al*., *Geophys. Res. Lett.* **24**, 2885–2888 (1997).
4. D. J. McComas *et al*., *J. Geophys. Res.* **105**, 10419–10433 (2000).
5. S. R. Cranmer, *Living Rev. Sol. Phys.* **6**, 3 (2009).
6. Y.-M. Wang, Y.-K. Ko, *Astrophys. J.* **880**, 146 (2019).
7. S. D. Bale *et al*., *Nature* **576**, 237–242 (2019).
8. G. Poletto, *Living Rev. Sol. Phys.* **12**, 7 (2015).
9. N. M. Viall, J. E. Borovsky, *J. Geophys. Res. Space Phys.* **125**, JA026005 (2020).
10. S. R. Cranmer, A. A. van Ballegooijen, R. J. Edgar, *Astrophys. J. Suppl. Ser.* **171**, 520–551 (2007).
11. T. Matsumoto, T. K. Suzuki, *Astrophys. J.* **749**, 8 (2012).
12. M. Shoda, B. D. G. Chandran, S. R. Cranmer, *Astrophys. J.* **915**, 52 (2021).
13. C.-Y. Tu *et al*., *Science* **308**, 519–523 (2005).
14. L. Yang *et al*., *Astrophys. J.* **770**, 6 (2013).
15. Y.-M. Wang, *Astrophys. J.* **904**, 199 (2020).
16. H. Tian *et al*., *Science* **346**, 1255711 (2014).
17. P. Kayshap, K. Murawski, A. K. Srivastava, B. N. Dwivedi, *Astron. Astrophys.* **616**, A99 (2018).
18. J. Gorman, L. P. Chitta, H. Peter, *Astron. Astrophys.* **660**, A116 (2022).
19. H. Tian, S. W. McIntosh, S. R. Habbal, J. He, *Astrophys. J.* **736**, 130 (2011).
20. S. Pucci, G. Poletto, A. C. Sterling, M. Romoli, *Astrophys. J.* **793**, 86 (2014).
21. N.-E. Raouafi, G. Stenborg, *Astrophys. J.* **787**, 118 (2014).
22. V. M. Uritsky *et al*., *Astrophys. J.* **907**, 1 (2021).
23. P. Kumar *et al*., *Astrophys. J.* **933**, 21 (2022).
24. N. E. Raouafi *et al*., *Astrophys. J.* **945**, 28 (2023).
25. I. A. Ahmad, G. L. Withbroe, *Sol. Phys.* **53**, 397–408 (1977).
26. Y.-M. Wang, *Astrophys. J.* **435**, L153 (1994).
27. S. Patsourakos, J.-C. Vial, *Astron. Astrophys.* **359**, L1–L4 (2000).
28. L. Teriaca, G. Poletto, M. Romoli, D. A. Biesecker, *Astrophys. J.* **588**, 566–577 (2003).
29. N. Fargette *et al*., *Astrophys. J.* **919**, 96 (2021).
30. P. Rochus *et al*., *Astron. Astrophys.* **642**, A8 (2020).
31. D. Müller *et al*., *Astron. Astrophys.* **642**, A1 (2020).
32. Materials and methods are available as supplementary materials.
33. F. Moreno-Insertis, K. Galsgaard, I. Ugarte-Urra, *Astrophys. J.* **673**, L211–L214 (2008).
34. K. Shibata *et al*., *Science* **318**, 1591–1594 (2007).
35. A. C. Sterling, R. L. Moore, D. A. Falconer, M. Adams, *Nature* **523**, 437–440 (2015).
36. S. Mandal *et al*., *Astron. Astrophys.* **664**, A28 (2022).
37. L. P. Chitta, A. R. C. Sukarmadji, L. Rouppe van der Voort, H. Peter, *Astron. Astrophys.* **623**, A176 (2019).
38. L. D. Xia, E. Marsch, W. Curdt, *Astron. Astrophys.* **399**, L5–L9 (2003).
39. G. L. Withbroe, *Astrophys. J.* **325**, 442 (1988).
40. E. N. Parker, *Astrophys. J.* **330**, 474 (1988).
41. K. L. Harvey, F. Recely, *Sol. Phys.* **211**, 31–52 (2002).
42. T. Sakao *et al*., *Science* **318**, 1585–1588 (2007).
43. G. A. Doschek *et al*., *Astrophys. J.* **667**, L109–L112 (2007).
44. L. K. Harra *et al*., *Astrophys. J.* **676**, L147–L150 (2008).
45. D. H. Brooks *et al*., *Astrophys. J.* **917**, 25 (2021).
46. Y.-M. Wang, in *Cool Stars, Stellar Systems, and the Sun*, R. A. Donahue, J. A. Bookbinder, Eds., vol. 154 of *Astronomical Society of the Pacific Conference Series* (1998), pp. 131–152.
47. S. Parhi, S. T. Suess, M. Sulkanen, *J. Geophys. Res.* **104**, 14781–14787 (1999).
48. J. Andries, M. Goossens, *Astron. Astrophys.* **368**, 1083–1094 (2001).
49. Th. Roudier, R. Muller, *Sol. Phys.* **107**, 11–26 (1986).
50. G. W. Simon, N. O. Weiss, *Z. Astrophys.* **69**, 435–450 (1968).
51. B. Mampaey *et al*., SolO/EUI Data Release 5.0 2022-04 (Royal Observatory of Belgium, 2022); https://doi.org/10.24414/2qfw-tr95.
52. I. Zouganelis *et al*., *Astron. Astrophys.* **642**, A3 (2020).

ACKNOWLEDGMENTS

Solar Orbiter is a space mission with international collaboration between ESA and NASA, operated by ESA. The EUI instrument was built by CSL, IAS, MPS, MSSL/UCL, PMOD/WRC, ROB, and LCF/IO with funding from the Belgian Federal Science Policy Office (BELSPO/PRODEX PEA 4000134088), the Centre National d'Etudes Spatiales (CNES), the UK Space Agency (UKSA), the Bundesministerium für Wirtschaft und Energie (BMWi) through the Deutsches Zentrum für Luft- und Raumfahrt (DLR), and the Swiss Space Office (SSO). This research has made use of NASA's Astrophysics Data System. **Funding:** L.P.C. acknowledges funding by the European Union (ERC, ORIGIN, 101039844). S.P. acknowledges funding by CNES through the Multi Experiment Data & Operation Center (MEDOC). D.M.L. thanks the Science and Technology Facilities Council for the award of an Ernest Rutherford Fellowship (ST/R003246/1). A.N.Z., D.B., E.K., L.R., and C.V. thank the Belgian Federal Science Policy Office (BELSPO) for the provision of financial support in the framework of the PRODEX Programme of the European Space Agency (ESA) under contract nos. 4000134474 and 4000136424. **Author contributions:** L.P.C. led the study and data analysis and wrote the manuscript with inputs from A.N.Z., D.B., H.P., S.P., S.M., R.A.C., U.S., L.T., É.B., D.M.L., and D.B.S. F.A., E.K., and C.V. contributed to data reduction. A.N.Z. led the Solar Orbiter observing campaign. D.B. is the principal investigator of EUI. All authors discussed and interpreted the results. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** The HRI_{EUV} level-2 data are archived by the Royal Observatory of Belgium (*51*). We used observations in the time range of 30 March 2022 04:30 to 05:00 UT (files from solo_L2_eui-hrieuvnon-image_20220330T043000227_V01.fits to solo_L2_eui-hrieuvnon-image_20220330T045957224_V01.fits). The data can alternatively be retrieved from ESA's Solar Orbiter Archive https://soar.esac.esa.int/soar/ using the same time range and the Solar Orbiter Observing Plan (SOOP) (*52*) name "R_SMALL_HRES_MCAD_Polar-Observations."
License information: Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse. This research was funded in whole or in part by the European Union through Horizon Europe (grant no. 101039844), a cOAlition S organization. The author will make the Author Accepted Manuscript (AAM) version available under a CC BY public copyright license.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.ade5801
Materials and Methods
Supplementary Text
Figs. S1 to S9
Tables S1 to S4
References (*53*–*75*)
Movies S1 to S12

Submitted 25 August 2022; accepted 14 July 2023
10.1126/science.ade5801

CARBON OFFSETS

Action needed to make carbon offsets from forest conservation work for climate change mitigation

Thales A. P. West[1,2]*, Sven Wunder[3,4], Erin O. Sills[5], Jan Börner[6,7], Sami W. Rifai[8], Alexandra N. Neidermeier[1], Gabriel P. Frey[6], Andreas Kontoleon[2,9]

Carbon offsets from voluntary avoided-deforestation projects are generated on the basis of performance in relation to ex ante deforestation baselines. We examined the effects of 26 such project sites in six countries on three continents using synthetic control methods for causal inference. We found that most projects have not significantly reduced deforestation. For projects that did, reductions were substantially lower than claimed. This reflects differences between the project ex ante baselines and ex post counterfactuals according to observed deforestation in control areas. Methodologies used to construct deforestation baselines for carbon offset interventions need urgent revisions to correctly attribute reduced deforestation to the projects, thus maintaining both incentives for forest conservation and the integrity of global carbon accounting.

For nearly two decades, the performance-based payment mechanism for reduced carbon emissions from deforestation and forest degradation known as REDD+ has been under intense debate (*1*). Although regulations and capacity for national REDD+ programs are still under development (*2*, *3*), many standalone, voluntary REDD+ projects are operational worldwide (*4*). These projects intend to conserve forests through many activities, such as improved monitoring and enforcement, promotion of sustainable practices, and local stakeholder engagement, often funded by the commercialization of carbon offsets [each corresponding to 1 Mg of carbon dioxide (CO_2) either removed from or not emitted into the atmosphere]. In 2021, two-thirds of the 227.7 million offsets from the land-use sector (excluding agriculture) traded in carbon markets, with a total value of USD $1.3 billion, originated from REDD+ projects (*5*).

Numerous policy discussions and initiatives focus on how to scale and integrate the carbon-emission reductions claimed by voluntary carbon-offset projects, particularly from REDD+ activities, into climate policies and Nationally Determined Contributions (NDCs) reported to the United Nations Framework Convention on Climate Change (*3*, *6*–*8*). However, there is little rigorous evidence on the contributions of these projects (*9*, *10*), with some studies suggesting that many are associated with little or no actual emission reductions (*11*–*17*).

Carbon offsets from REDD+ projects are issued on the basis of comparison between the observed forest cover in the project areas and deforestation baseline scenarios expected to have been realized in the absence of REDD+, which are de facto unobservable (*3*, *17*). Many project baselines are formed through the extrapolation of historical deforestation averages or trends, often spatially projected over a reference region that encompasses the project sites (*17*). These crediting baselines may become unrealistic counterfactuals with extensive changes in economic or political conditions known to affect deforestation rates (*18*), combined with questionable modeling decisions underlying the spatial projections (*19*, *20*). Baselines could also be opportunistically inflated by profiteers seeking to maximize the volume of offsets issued by a project (*21*). As a result, carbon offsets may lack "additionality"—they may not reflect actual emission reductions (*22*).

This study provides a pantropical comparison between ex post deforestation counterfactuals, informed by observable control areas, and the ex ante baselines adopted by 27 voluntary REDD+ projects in six tropical countries: Peru, Colombia, Democratic Republic of Congo (DRC), Tanzania, Zambia, and Cambodia (Fig. 1, figs. S1 and S2, and tables S1 and S2) certified under the Verified Carbon Standard (*23*). Because some projects are composed of multiple disconnected sites, we evaluated those individually, increasing our sample to 31 project sites. We present both project-specific and cross-project

[1]Environmental Geography Group, Institute for Environmental Studies (IVM), VU University Amsterdam, Amsterdam, Netherlands. [2]Centre for Environment, Energy and Natural Resource Governance, University of Cambridge, Cambridge, UK. [3]European Forest Institute (EFI), Barcelona, Spain. [4]Center for International Forestry Research (CIFOR), Lima, Peru. [5]Department of Forestry and Environmental Resources, North Carolina State University, Raleigh, USA. [6]Center for Development Research (ZEF), University of Bonn, Bonn, Germany. [7]Institute for Food and Resource Economics (ILR), University of Bonn, Bonn, Germany. [8]ARC Centre of Excellence for Climate Extremes, University of New South Wales, Sydney, Australia. [9]Department of Land Economy, University of Cambridge, Cambridge, UK.
*Corresponding author. Email: t.a.pupowest@vu.nl

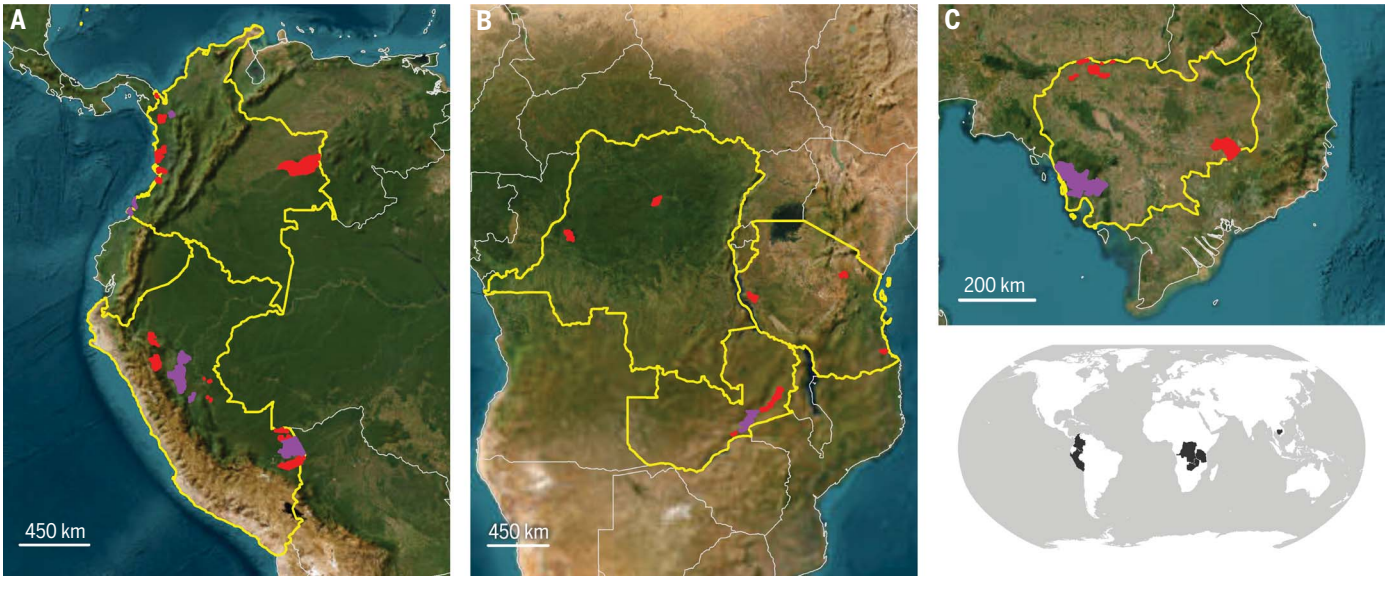


Fig. 1. Voluntary REDD+ project sites included in the study. (**A**) Peru and Colombia. (**B**) Democratic Republic of Congo (DRC), Tanzania, and Zambia. (**C**) Cambodia. Study areas are indicated in red. Purple areas are the sites excluded from the analysis.

analyses based, respectively, on the standard and generalized versions of the synthetic control (SC) method for causal inference (*24*, *25*) that estimate the reductions in deforestation in project sites attributable to the REDD+ interventions.

The SCs are constructed to be based on control areas ("donors"), selected from project-specific donor pools, which had similar levels of forest cover and other characteristics (table S2) and were exposed to similar levels of deforestation pressure (as determined by comparing the average annual deforestation in the projects' and donors' 1-km and 10-km buffer zones before the project implementation). The SC method selects and constructs a weighted combination of control areas that has similar characteristics as those of the REDD+ site and follows a similar historical trajectory of deforestation (supplementary materials). Before interpreting individual project results, we conducted project-specific "validation" tests to check whether the standard SC method (*24*) was able to construct SCs with deforestation rates similar to that of project areas during the immediate preproject period (*17*). Conservatively, we focus the discussion of our results on projects with SCs that performed well in the validation test (fig. S5 and table S3).

Average impacts were estimated with the generalized SC (GSC) method (*25*) on the basis of two independent sets of control areas. Furthermore, to address the concern that the selected control areas may not represent potential counterfactuals for REDD+ project sites, we also estimated average project impacts through comparison of operational project sites with "yet-to-become" project areas throughout the study period using matching-based methods for time-series cross-sectional analyses (*26*). Because the evaluated projects span multiple countries and contexts, our analyses shed light on the robustness of the assumptions adopted for the construction of REDD+ baselines under a wide range of deforestation conditions (fig. S2).

Individual REDD+ project impacts

The individual SC analyses show mixed impacts of the voluntary REDD+ projects on deforestation. Results from the validation tests suggest that the SC method could replicate pre-REDD+ deforestation trends in the "to-be" project sites (fig. S6 and table S3). We discarded only one project (1775-1) from the analyses because of the poor fit between the deforestation in the SC and the REDD+ site before project implementation in our validation test. Four other projects (985, 1360-1, 1389, and 1748) were also discarded out of an abundance of caution because despite the agreement in preproject deforestation between the SCs and the REDD+ sites, their buffer deforestation rates suggested substantially different levels of deforestation pressure. Our final sample was thus reduced to 26 project sites.

Eight of the remaining 26 project sites showed some evidence of additional reductions in deforestation compared with their individual SCs (figs. S7 and S8), although generally not to the extent claimed by the projects based on their ex ante crediting baselines. Additionality was most likely in Peru, where half of the REDD+ sites had significantly less deforestation than that of the ex post counterfactuals, with statistical significance judged by means of placebo tests. Three of the seven Colombian project sites and one of two Cambodian sites achieved significant deforestation reductions according to the SCs and placebo tests. No evidence of avoided deforestation was found for the REDD+ sites in the DRC, Tanzania, and Zambia with regard to their counterfactuals.

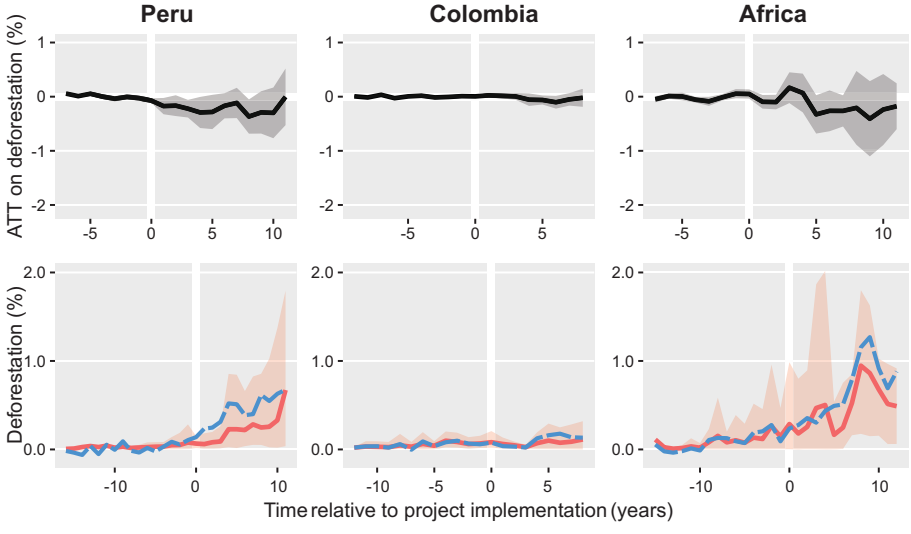


Fig. 2. Estimated average impacts of REDD+ projects in Peru, Colombia, and Africa on annual deforestation. Averages are based on the GSC method and the donor pool of control areas selected for the individual project's SCs. (**Top**) The average treatment effect on the treated (ATT) project sites. (**Bottom**) Projects' (solid red line) and counterfactuals' (dashed blue line) deforestation averages. Shaded red areas indicate bootstrapped 95% confidence intervals around the projects' deforestation average.

Average REDD+ project impacts

Average project impacts on deforestation [average treatment effects on the treated (ATT)] in Peru, Colombia, and Africa (DRC, Tanzania, and Zambia) were estimated with the GSC method (Fig. 2, top). Cambodian projects were excluded from this analysis because of the limited sample size. Unlike the individual project evaluations, the GSC analyses were based exclusively on annual deforestation rates and time-variant covariates. The GSC analyses were based on two independent sets of selected control areas for each region: In the first set, only the donors selected for the construction of the individual SCs were considered, whereas in the second set, controls were selected through cardinality matching (*27*), independent of the SC analyses. Our conclusions are robust to all of these methodological variations.

For the first set of controls, the average impact of the Peruvian projects on forest loss over 10 years was −0.24% or avoided deforestation of 686 ha year^{-1} (table S7). This effect was statistically significant in the first 4 years of project implementation (Fig. 2, bottom). An ATT of −0.14% or 414 ha year^{-1} (table S8) was found for the African projects, whereas a smaller effect was associated with the Colombian projects (−0.03% or 49 ha year^{-1}) (table S9). Neither the estimates for the Colombian nor the African groups of projects were statistically significant. Even assuming the estimated average reductions in deforestation to be significant in all three regions (a plausible assumption given our small sample sizes), they would still be substantially lower than avoided deforestation calculated from the average ex ante baseline deforestation rates adopted by the projects in Peru (3661 ha year^{-1}), Colombia (2550 ha year^{-1}), and Africa (2700 ha year^{-1}) through 2020.

These results are robust to using control areas selected through cardinality matching. On the basis of our second control set, we estimated ATTs of the Peruvian, Colombian, and African REDD+ sites as −0.42% (1266 ha) year^{-1}, −0.01% (30 ha) year^{-1}, and −0.21% (423 ha) year^{-1}, respectively (fig. S10, top, and tables S10 to S12). Again, only the estimate for the Peruvian projects was statistically significant, and only in the first 4 years of project implementation (fig. S10, bottom). Although the estimated impacts in Peru were −0.18 percentage points larger as compared with the first control set, the translated absolute average reduction still represents just one-third of the average reductions claimed by the Peruvian projects.

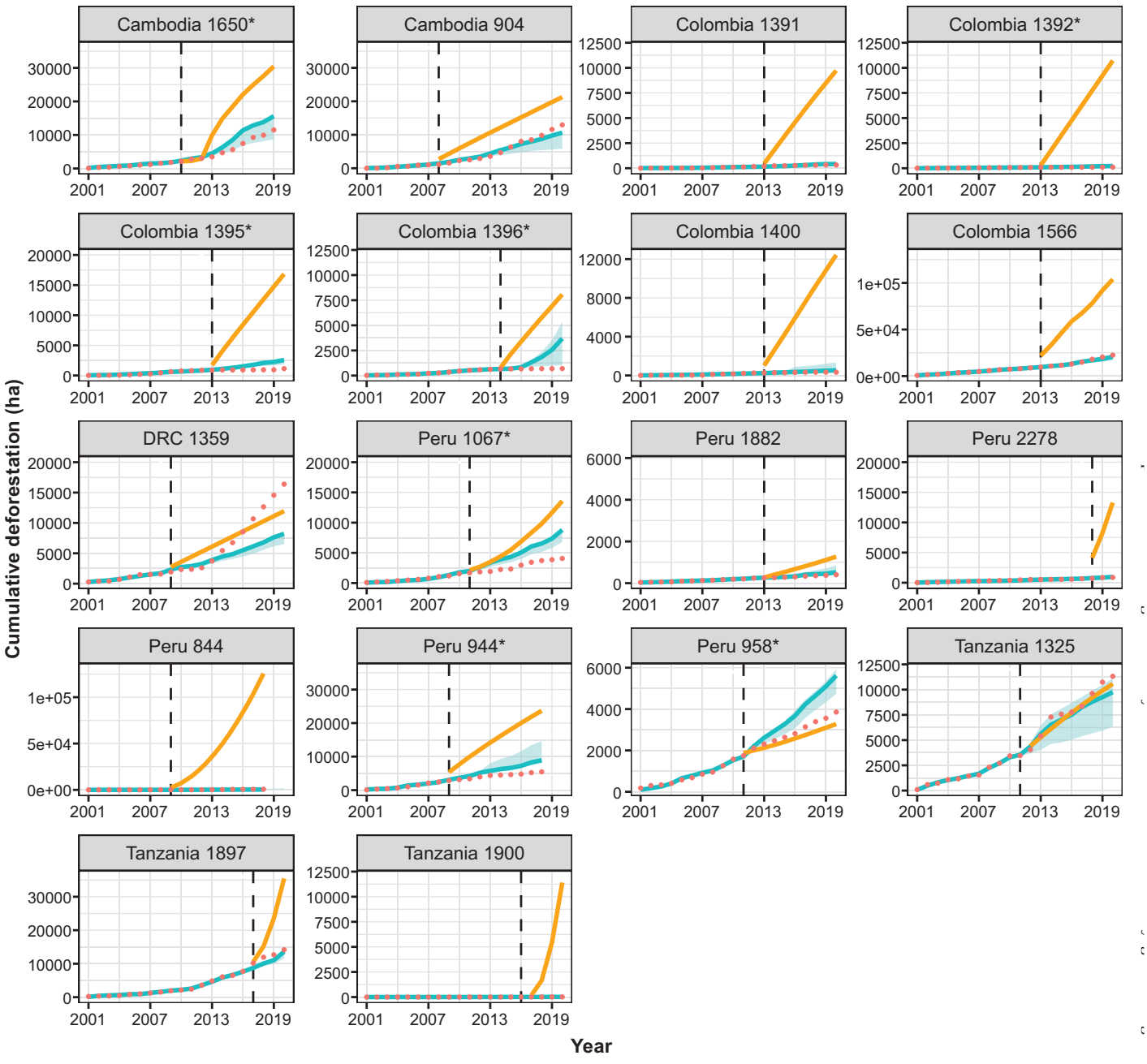


Fig. 3. Cumulative deforestation from the baseline scenarios adopted by the REDD+ projects versus observed cumulative deforestation in the SCs. Orange, REDD+ projects; SCs, blue. Shaded blue areas are possibility intervals. Dotted red lines indicate the observed cumulative deforestation in the project sites. Dashed black lines indicate the project implementation year. Asterisks indicate a significant reduction in deforestation in the REDD+ site compared with the SC based on placebo tests (scales differ).

Last, results from the analysis of already operational versus "to-become" project sites corroborate the findings from the GSC analyses. The average REDD+ impacts on deforestation ranged from −0.01% to 0.12% year^{-1} (or −113 to 121 ha year^{-1}), across different model specifications, but the estimates were not significant in any of the regions (fig. S14).

Carbon-offset implications

We investigated the implications of our findings for the environmental integrity of the credits issued by the REDD+ projects. These implications are based on the 18 out of 26 projects with sufficient publicly available information about baseline deforestation rates (Fig. 3 and tables S5 and S6). Only one project baseline was lower than its SC (project 958), and only one project baseline was similar to its SC (project 1325). All other ex ante baselines posited more deforestation than estimated ex post according to SCs. According to the projects' ex ante estimates, up to 89 million carbon offsets could potentially have been generated by these 18 REDD+ projects through 2020. Yet 60.2 million of these offsets (68%) would have originated from projects that have not significantly reduced deforestation (and carbon emissions) compared with their SCs. The remaining 28.8 million offsets (32%) would have originated from projects likely associated with some avoided deforestation, but not to the extent expected by the project developers. If we replace the ex ante baselines adopted by the projects with the deforestation observed in the SCs, our estimates suggest that only 5.4 million (6.1%) of the

89 million expected offsets from the REDD+ projects would be associated with additional carbon emission reductions.

As of November 2021, those 18 REDD+ projects had issued 62 million carbon-offset credits (table S4). Out of those, at least 14.6 million (24%) have already been used by individuals or organizations around the world to offset their greenhouse gas emissions. Thus, according to our SC-based estimates, these projects have already been used to offset almost three times more carbon emissions than their actual contributions to climate change mitigation—with another 47.4 million carbon offsets being readily available in the market.

Discussion

Overall, the weight of evidence suggests that the voluntary REDD+ projects in our sample across six tropical countries achieved much less avoided deforestation than forecast by project developers. Only a minority of projects achieved statistically significant reductions in comparison with ex post counterfactuals. Our findings corroborate prior studies that questioned the additionality, and thus environmental integrity, of carbon-offset interventions (*11*–*17*). Ex ante baselines that exaggerate the deforestation that would occur without REDD+, likely facilitated by methodological flexibility in their construction and exacerbated by adverse site selection (*14*), are a major reason for the gap between offsets projected ex ante and actual offsets estimated ex post. Poor performance by the REDD+ projects (failure to de facto reduce deforestation) may also be a factor (*21*, *28*).

In an evaluation of voluntary REDD+ projects in the Brazilian Amazon, West *et al.* (*17*) pointed to the potential confounding effect created by Brazil's post-2004 policy interventions to control deforestation, triggering a widespread reduction in forest loss between 2004 and 2012 (*29*). As a result, the high regional deforestation rates observed before 2004, used to inform the Brazilian project ex ante baselines, likely led to an overestimation of the projects' performance. Yet unlike Brazil, the six countries in this study did not experience a similar nationwide reduction in deforestation after the REDD+ projects were implemented (fig. S3). Hence, the unrealistic ex ante baselines adopted by many projects likely resulted from the use of methodologies that systematically fail to produce credible counterfactuals for the REDD+ interventions, compromising the evaluation of the projects' effect on mitigating deforestation and thus carbon emissions. This may be due to potentially four complementary reasons: poor foresight, adverse site selection, limited room for adjustments over time, and "gaming." First, projects may have (unintentionally) overestimated future deforestation pressures by following baseline methodologies that heavily rely on the continuation of historical trends that no longer represent current conditions, often combined with problematic spatial deforestation projections (*19*, *30*). Second, projects may have been preferentially sited where conditions are conducive to reducing deforestation and therefore where deforestation may have been lower than suggested by ex ante baselines constructed from deforestation trends in the surrounding region (*14*). Third, rules adopted by certification standards require baselines to be fixed for a period of usually 10 years, restricting adjustments to reflect changes in deforestation drivers over time (*23*). Last, flexibility in baseline methodologies may have been opportunistically exploited to maximize revenues from offset sales.

By contrast, both standard and generalized SC methods use pre-REDD+ information to identify control areas but use contemporaneous ex post information on deforestation in the project and control areas to measure additionality. Following well-accepted guidelines for rigorous impact evaluation (*31*), if properly selected, such ex post counterfactuals can capture the effects of contemporaneous changes in deforestation drivers and thus are less likely to generate effect estimates confounded by external factors (*10*, *17*). REDD+ projects adopting a similar dynamic approach would likely reduce the additionality problems with project ex ante baselines and offsets identified in this study.

Despite the clear advantages (from a causal inference perspective) of using ex post methods such as SCs to construct deforestation baselines for REDD+ interventions, some implementation and monitoring challenges would likely arise from their adoption. First, given the biophysical heterogeneity of tropical regions, and limited data, ideal control areas for the project sites may not always exist or be possible to identify (supplementary materials, appendix A). Second, those control areas could be manipulated (for example, intentionally degraded or conserved) to misleadingly improve or reduce estimated project performance. Third, dynamic baselines may still fail to account for all relevant determinants of deforestation owing to data constraints. Last, long-lasting voluntary REDD+ projects may eventually outlive their SCs as new interventions or other regime shifts occur in control sites.

One alternative would be to require projects to adopt transparent ex ante jurisdictional baselines that are preestablished by government agencies. Although these baselines might still fall short at demonstrating causal links and thus additionality, they could be both more transparent and updated more frequently than individual project baselines to reflect emerging deforestation pressures and spatial patterns. Reductions in deforestation relative to jurisdictional baselines would still have to be correctly attributed to either government efforts to control forest loss or to private REDD+ interventions (*10*), in addition to changes in external deforestation drivers (such as agricultural commodity prices). Transferring the responsibility of baseline construction from project developers to jurisdictions could reduce the room for "baseline gaming," although the risk of adverse site selection could remain (*14*). On balance, nesting voluntary projects into subnational jurisdictions appears to be a promising future pathway for REDD+, a practice that is gaining traction globally (*6*). However, the specifics of implementation are crucial and require careful consideration.

Another possible explanation for limited additionality is poor performance by the projects. This would be likely if projects were failing to control deforestation in project areas. We observed generally low deforestation rates in project areas, but we cannot compare that with the projects' targeted deforestation rates in those areas because those are not always reported in the publicly available information on projects. However, it is clear that some projects have struggled with on-the-ground implementation and execution of envisioned conservation activities; others may have promoted ineffective actions, whether because of funding uncertainties, slow commercialization of carbon credits, lack of experience, or poor management (*21*, *28*, *32*). Many projects claimed to have started much earlier than the year they were certified. Although this allows projects to issue and sell retroactive offsets immediately after certification (*33*), it also implies that they did not have access to carbon funding during their initial years, potentially compromising the execution of planned conservation actions.

A recent evaluation on the effectiveness of the same type of REDD+ interventions reported significant reductions in average deforestation rates (*34*). The study, based on satellite pixel matching, estimated an average deforestation rate of 0.2% $year^{-1}$ in the REDD+ sites versus 0.4% $year^{-1}$ for their matched control areas. The size of these estimates is similar to some of our estimates from GSC analyses; but in our case, they were statistically insignificant, potentially because of our smaller sample sizes compared with the pixel-based samples from the previous study. And from the offsetting perspective, the average reductions—significant or not—were substantially lower than the additional reductions in deforestation projected and claimed by the projects.

Our study provides further evidence on the effectiveness of voluntary REDD+ projects and questions their de facto additionality (*35*). Only a minority of the projects significantly reduced deforestation in the project areas compared with the ex post counterfactuals, and even those, with one exemption, did not reduce deforestation to the extent claimed. Although REDD+ payments are typically conditioned on performance in project areas, only the offsets

associated with additional reductions in deforestation relative to a counterfactual genuinely offset emissions of potential buyers in the voluntary carbon markets (*7*, *21*). Certification schemes are allegedly in place to safeguard the additionality of offsets, but our results indicate that currently used baseline methodologies do not guarantee additionality. It is critical to develop new and rigorous methods for the construction of credible deforestation baselines for voluntary REDD+ interventions and to properly and regularly assess their contribution to climate change mitigation.

Last, the evidence from this and other studies indicates that some voluntary projects have effectively reduced deforestation (*34*), particularly in Peru. For REDD+ to be scaled and achieve its ambitious goals worldwide, it is paramount that we better understand the factors that drive both mitigation performance and impacts on local communities. Building on this knowledge, academics, practitioners, and policy-makers must form effective partnerships to help REDD+ fulfill its original promise.

REFERENCES AND NOTES

1. A. Angelsen, *Rev. Dev. Econ.* **21**, 237–264 (2017).
2. J. Börner *et al*., in *Transforming REDD+: Lessons and new directions*, A. Angelsen *et al*., Eds. (CIFOR, 2018), pp. 105–116.
3. FAO, "From reference levels to results reporting: REDD+ under the United Nations Framework Convention on Climate Change. 2019 update" (Food and Agriculture Organization, Rome, 2019).
4. S. S. Atmadja *et al*., *Environ. Res. Lett.* **17**, 044038 (2022).
5. S. Donofrio, P. Maguire, C. Daley, C. Calderon, K. Lin, *The Art of Integrity: State of the Voluntary Carbon Markets 2022 Q3* (Forest Trends' Ecosystem Marketplace, 2022).
6. D. Lee, P. Llopis, R. Waterworth, G. Roberts, T. Pearson, *Approaches to REDD+ Nesting: Lessons Learned from Country Experiences* (World Bank, 2018).
7. Taskforce on Scaling Voluntary Carbon Markets, *Phase 1—Final Report* (Institute of International Finance, 2021).
8. Voluntary Carbon Markets Integrity Initiative, *Aligning Voluntary Carbon Markets with the 1.5°C Paris Agreement Ambition* (Voluntary Carbon Markets Integrity Initiative, 2021).
9. A. E. Duchelle, G. Simonet, W. D. Sunderlin, S. Wunder, *Curr. Opin. Environ. Sustain.* **32**, 134–140 (2018).
10. E. O. Sills *et al*., *Glob. Environ. Change* **43**, 148–160 (2017).
11. G. Badgley *et al*., *Glob. Change Biol.* **28**, 1433–1445 (2022).
12. R. Calel, J. Colmer, A. Dechezleprêtre, M. Glachant, *SSRN Electron. J.* **5709** (2021).
13. M. Cames *et al*., *How Additional is the Clean Development Mechanism?* (Öko-Institut e.V., 2016).
14. P. Delacote, G. Le Velly, G. Simonet, *Resour. Energy Econ.* **67**, 101277 (2022).
15. B. Haya *et al*., *Clim. Policy* **20**, 1112–1126 (2020).
16. A. Kollmuss, L. Schneider, V. Zhezherin, *Has Joint Implementation Reduced GHG Emissions? Lessons Learned for the Design of Carbon Market Mechanisms* (Stockholm Environment Institute, 2015).
17. T. A. P. West, J. Börner, E. O. Sills, A. Kontoleon, *Proc. Natl. Acad. Sci. U.S.A.* **117**, 24188–24194 (2020).
18. J. Busch, K. Ferretti-Gallon, *Rev. Environ. Econ. Policy* **11**, 3–23 (2017).
19. S. Sloan, J. Pelletier, *Glob. Environ. Change* **22**, 440–453 (2012).
20. M. Ehara *et al*., *For. Policy Econ.* **129**, 102474 (2021).
21. C. Seyller *et al*., *Int. For. Rev.* **18**, 231–246 (2016).
22. S. Wunder, *Conserv. Biol.* **21**, 48–58 (2007).
23. Verra, VCS Standard, version 4.0 (Verra, 2019).
24. A. Abadie, A. Diamond, J. Hainmueller, *J. Stat. Softw.* **42**, 1–17 (2011).
25. Y. Xu, *Polit. Anal.* **25**, 57–76 (2017).
26. K. Imai, I. S. Kim, E. H. Wang, *Am. J. Pol. Sci.* **0**, 1–19 (2021).
27. J. R. Zubizarreta, R. D. Paredes, P. R. Rosenbaum, *Ann. Appl. Stat.* **8**, 204–231 (2014).
28. M. S. Verra, Suruí Forest Carbon Project (2018); https://verra.org/media-statement-surui-forest-carbon-project.
29. T. A. P. West, P. M. Fearnside, *Land Use Policy* **100**, 105072 (2021).
30. O. Mertz *et al*., *J. Land Use Sci.* **13**, 1–15 (2018).
31. P. J. Ferraro, M. M. Hanauer, *Annu. Rev. Environ. Resour.* **39**, 495–517 (2014).
32. S. Wunder *et al*., *Front. For. Glob. Change* **3**, 11 (2020).
33. N. Linacre, R. O'Sullivan, D. Ross, L. Durschinger, I. Deshmukh, "REDD+ Supply and Demand 2015–2025" (US Agency for International Development, 2015).
34. A. Guizar-Coutiño, J. P. G. Jones, A. Balmford, R. Carmenta, D. A. Coomes, *Conserv. Biol.* **36**, e13970 (2022).
35. A. Angelsen, *Moving ahead with REDD: Issues, Options and Implications* (CIFOR, 2008), vol. 6.
36. T. A. P. West, Replication Data for: Action needed to make carbon offsets from forest conservation work for climate change mitigation. *DataverseNL* (2023); https://doi.org/10.34894/IQC9LM.

ACKNOWLEDGMENTS

We thank the German Development Agency (GIZ) for their support during the planning phase of this study. **Funding:** This research was supported by Norway's International Climate and Forest Initiative (NICFI), the Meridian Institute, CIFOR's Global Comparative Study on REDD+, and the European Forest Institute's BMEL-financed NewGo project. J.B. acknowledges partial funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy, EXC 2070–390732324. **Author contributions:** Conceptualization: T.A.P.W., S.W., E.O.S., J.B., and A.K. Methodology: T.A.P.W., S.W., E.O.S., J.B., and A.K. Data processing: T.A.P.W., S.W.R., A.N.N., and G.F.; Formal analyses: T.A.P.W. Visualization: T.A.P.W. Writing: T.A.P.W., S.W., E.O.S., J.B., S.W.R., A.N.N., and A.K. Funding acquisition: T.A.P.W, S.W., and J.B. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** The data and codes used in this study can be accessed through the DataverseNL repository (*36*). **License information:** Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.ade3535
Materials and Methods
Figs. S1 to S14
Tables S1 to S12
Appendix A
References (*37*–*46*)

Submitted 15 August 2022; accepted 7 July 2023
10.1126/science.ade3535

ORGANIC CHEMISTRY

Aminodealkenylation: Ozonolysis and copper catalysis convert C(sp^3)–C(sp^2) bonds to C(sp^3)–N bonds

Zhiqi He, Jose Antonio Moreno, Manisha Swain, Jason Wu, Ohyun Kwon*

Great efforts have been directed toward alkene π bond amination. In contrast, analogous functionalization of the adjacent C(sp^3)–C(sp^2) σ bonds is much rarer. Here we report how ozonolysis and copper catalysis under mild reaction conditions enable alkene C(sp^3)–C(sp^2) σ bond–rupturing cross-coupling reactions for the construction of new C(sp^3)–N bonds. We have used this unconventional transformation for late-stage modification of hormones, pharmaceutical reagents, peptides, and nucleosides. Furthermore, we have coupled abundantly available terpenes and terpenoids with nitrogen nucleophiles to access artificial terpenoid alkaloids and complex chiral amines. In addition, we applied a commodity chemical, α-methylstyrene, as a methylation reagent to prepare methylated nucleosides directly from canonical nucleosides in one synthetic step. Our mechanistic investigation implicates an unusual copper ion pair cooperative process.

The importance of aliphatic amines and nitrogen heterocycles is evidenced by their wide representation in natural products, pharmaceuticals, agrochemicals, and other bioactive compounds (*1*). The increasing demand in medicinal research for more-complex three-dimensional architectures and optically active amines has driven efforts to develop diverse practical methods for C(sp^3)–N bond construction (*2*, *3*). Conventionally, the electron-rich nature of amino nitrogen atoms renders them strong nucleophiles that can attack the polarized chemical bonds of alkyl halides, alcohol derivatives, and carbonyls, thereby forging new C(sp^3)–N bonds. Engaging ubiquitous and nonpolar C–C bonds in C(sp^3)–N bond construction is an intriguing strategy because such transformations can be used to modify molecular skeletons directly and, thereby, access conventionally challenging or inaccessible aliphatic amines (*4*–*6*). Classical methods for C–C bond amination, including the Lossen, Hofmann, Curtius, Beckmann, and Schmidt rearrangements, which typically rely on 1,2-migration of an alkyl or aryl group from carbon to nitrogen (Fig. 1A, reactions 1 and 2), have demonstrated their utility in the syntheses of bioactive molecules and in industrial manufacturing. For example, 5.5 million metric tons of caprolactam, a precursor to nylon 6, are produced annually through the Beckmann rearrangement of cyclohexanone oxime (*7*). Although these classical transformations are powerful methods for accessing nitrogen insertion products, the sources of the C–C bonds are limited mostly to ketones and carboxylic acids, and the sources of nitrogen atoms are limited to those that insert one N atom, leaving room for incorporation of alternative, perhaps more ubiquitous, chemical functionalities as the sources of both the C–C and C–N bonds.

Department of Chemistry and Biochemistry, University of California, Los Angeles, CA 90095-1569, USA.
*Corresponding author. Email: ohyun@chem.ucla.edu

In the past 5 years, several important advances have sought to address those limitations in the sources of both the C–C bonds and nitrogen atoms in C–C amination. The Jiao group employed the C–C bonds of alkylarenes, styrenes, and alkynylarenes in a Schmidt-type reaction, although the products were limited to anilines with new C(sp^2)–N bonds, owing to the preferred migration of the aryl group (Fig. 1A, reaction 3) (*8*–*10*). Fu, Hu, MacMillan, Martin, and colleagues disclosed deconstructive C(sp^3)–N couplings of carboxylic acid and ketone derivatives that overcame the limitation of 1,2-migration chemistry, providing methods to construct C(sp^3)–N bonds with various N-heteroarenes, amides, and anilines (Fig. 1A, reaction 4) (*11*–*15*). Despite these advances, there remains a huge scope for new sources of C–C bonds, beyond carbonyl compounds, for amination (*16*–*20*). In particular, developing methods to exploit a starting material's skeletal complexity to access chiral amine architectures would be highly desirable.

Alkenes are versatile functional groups that are abundant in natural products and industrial chemicals. In fact, C=C double bonds are the second most frequently encountered functional group in natural products (39.9%), exceeding ketone (15.9%) and carboxylic acid (10.6%) functionalities (*21*). Because chiral centers are common within natural products, we surmised that the alkene moieties in natural products might serve as ideal C–C bond amination precursors for constructing complex chiral amines from naturally occurring chiral-pool molecules. Most methods for installing nitrogen groups to alkene motifs have focused on addition across the C–C π bond to produce peripheral amination products (Fig. 1B) (*3*). Conversely, the adjacent C(sp^3)–C(sp^2) σ bonds have received less attention for C–N bond construction (Fig. 1C, top). We envisioned that a C–N bond coupling strategy involving unusual C–C bond disconnections would constitute a distinct paradigm for alkyl amine synthesis, with several implications. First, this strategy could access artificial terpenoid alkaloids and complex chiral amine architectures by embedding nitrogen motifs into the natural product precursors (Fig. 1D, top left) (*22*). Second, a commodity chemical, α-methylstyrene, could be applied as an N-methylation reagent (Fig. 1D, top right; for comparison with other N-methylation reagents, see section 4.5 of the supplementary materials) (*23*). Third, such a tool could facilitate the generation of useful value-added compounds that have been challenging to form or inaccessible from readily available starting materials. One such example is (1*R*,2*R*,5*R*)-5-amino-2-methylcyclohexan-1-ol hydrogen chloride, a precursor of a c-Jun N-terminal kinase (JNK) inhibitor. The considerable step economy and cost savings from previous approaches indicate the potential utility of this C(sp^3)–C(sp^2) amination strategy in bioactive reagent synthesis (Fig. 1D, bottom) (*24*).

To accomplish the conversion of alkene C(sp^3)–C(sp^2) σ bonds to C(sp^3)–N bonds, we had to address a fundamental challenge: cleavage of a stronger alkene C(sp^3)–C(sp^2) σ bond [bond dissociation energy (BDE) = 102 kcal/mol for propene] preferably over an aliphatic alkyl C(sp^3)–C(sp^3) σ bond (90 kcal/mol in ethane) (*25*). To realize this, our strategy engaged ozone, widely applied in several industries, to activate alkenes and facilitate their C(sp^3)–C(sp^2) σ bond cleavages (*26*, *27*). Ozonolysis of alkenes in alcoholic solvents affords α-alkoxyhydroperoxides via the Criegee intermediate (*28*). The weakness of the O–O bond (44 to 46 kcal/mol for alkyl hydroperoxides) of an α-alkoxyhydroperoxide can serve as the energetic driving force to promote fragmentation of the robust C–C bond. Previously, we had used Fe(II) salts as single-electron reducing agents to instigate O–O scission. The resulting α-alkoxyalkoxyl radical undergoes β-scission to release an alkyl radical that can be captured by organic radicophiles to provide the functionalization products of seemingly inert C(sp^3)–C(sp^2) σ bonds (Fig. 1C, bottom) (*29*–*31*).

Copper is a redox-active metal that has been used for single-electron reduction of peroxide intermediates in both biological processes and synthetic chemistry (*32*). Recent studies have demonstrated that copper catalysts are efficient at trapping alkyl radicals to afford C(sp^3)–N bonds as reductive elimination products (*12*, *13*, *33*–*36*). We envisioned sequential reduction of the hydroperoxide intermediate by a copper(I) amido complex and trapping of the resultant alkyl radical by the copper(II) center to induce reductive elimination and, thereby, form the C(sp^3)–N bond in an overall redox-neutral process (Fig. 1C, bottom).

Reaction development

To implement our design, we investigated the impact of the copper salt, ligand, and solvent.

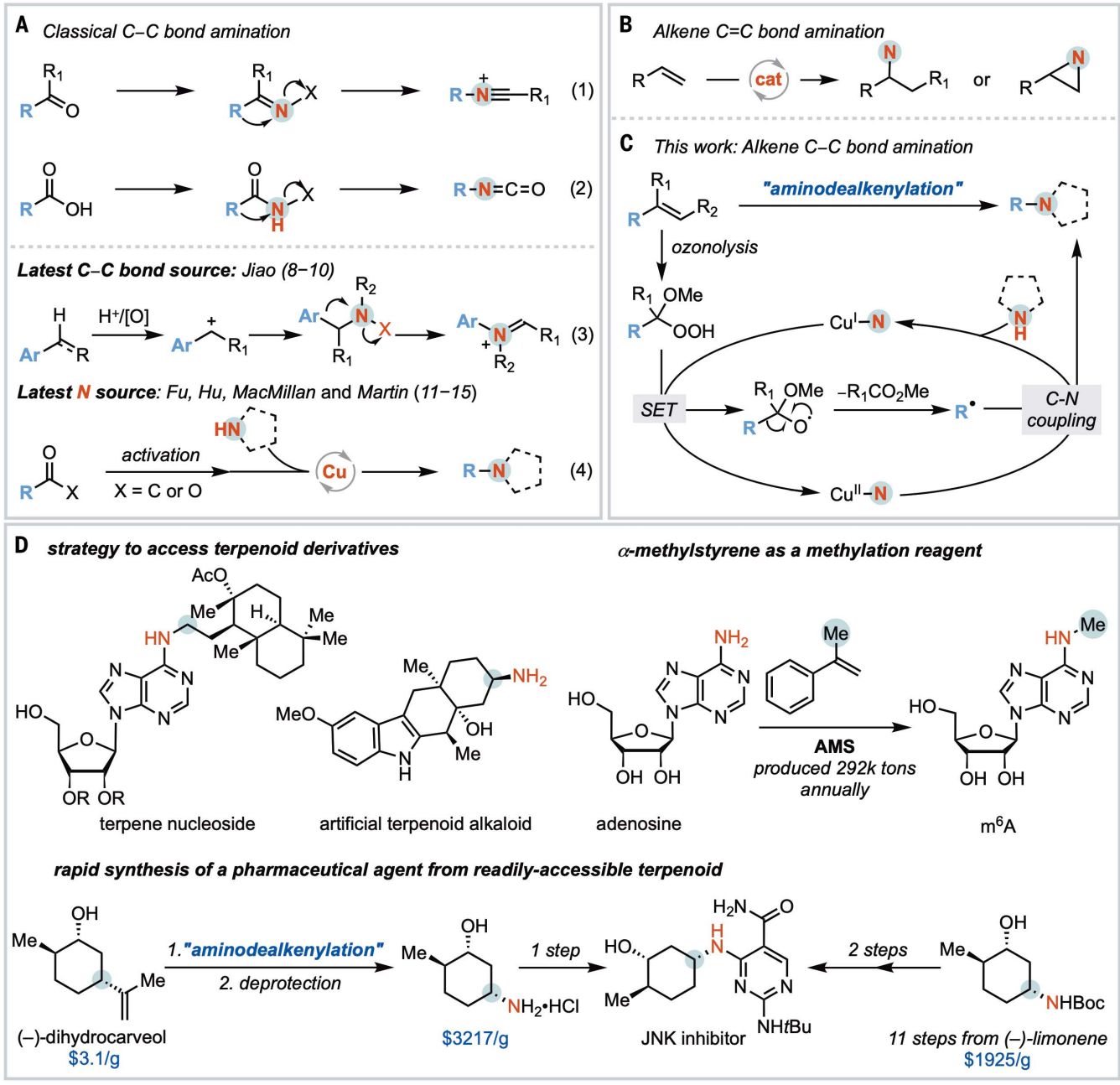


Fig. 1. Concept and development of dealkenylative C(sp^3)–N bond coupling. (**A**) Classic strategies in C(sp^3)–N bond construction from carbonyl compounds and recent advances. (**B**) C(sp^3)–N formation from alkene π bonds. (**C**) C(sp^3)–N formation from alkene σ bonds. (**D**) Applications of dealkenylative C(sp^3)–N bond coupling. SET, single-electron transfer; cat, catalysis; AcO, acetoxy; Me, methyl; *t*Bu, *tert*-butyl; Boc, *tert*-butoxycarbonyl.

We discovered that the aminodealkenylation could be performed efficiently under mild reaction conditions when using CuCl (20 mol %) and 1,10-phenathroline (20 mol %) in acetonitrile (MeCN) at room temperature (see the supplementary materials for details). After determining the reaction conditions, we assessed a diverse array of nitrogen nucleophiles for dealkenylative C–N coupling with (–)-isopulegol as the model alkene substrate (Fig. 2). Our protocol could be used to install functionalized indoles (**1** to **7**) into (–)-isopulegol in high yields, producing enantiomerically pure indole derivatives that are promising scaffolds for drug development (*37*). Two indole-based natural products, the sleep hormone melatonin (**6**) and a protected form of the amino acid tryptophan (**7**), were alkylated in high yields with exclusive regioselectivity in the presence of secondary amide N–H bonds. Other pharmaceutically important substituted heterocyclic compounds, including azaindoles (**8** to **10**), indazoles (**11** and **12**), carbazole (**13**), pyrazoles (**14** and **15**), and pyrrole (**16**), were successfully coupled with (–)-isopulegol in good yields. 6-Chloro-7-iodo-7-deazapurine (**17**), a privileged scaffold in antitumor and antiviral drugs, was also a competent alkylation substrate. We obtained high yields when coupling it with phthalimide (**18**), and this process could be scaled (50 mmol) with a slightly decreased yield. In addition to azacycles, we found that 2-aminopyrimidine (**19**), substituted anilines (**20** and **21**), and 2-aminopyridine (**22**) also underwent the dealkenylative C–N coupling efficiently. Less-nucleophilic amide (**23**) and sulfonamide (**24**) substrates also provided their

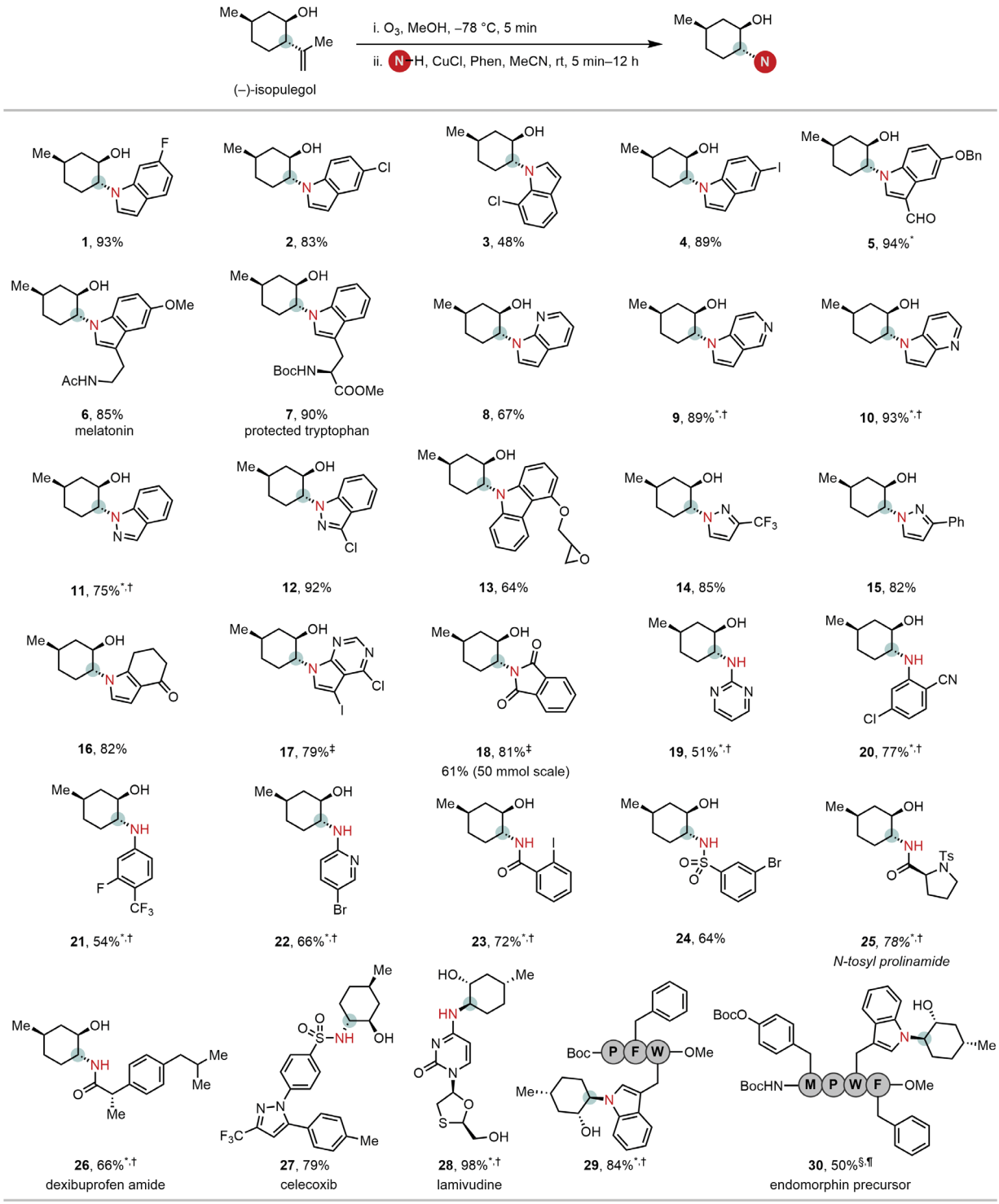


Fig. 2. Substrate scope of nitrogen nucleophiles. Reactions were performed on a 0.2- to 2.0-mmol scale using 20 mol % CuCl and 20 mol % phenanthroline (Phen) at 23°C in MeCN. All yields are isolated yields. Absolute configurations are indicated by wedged and dashed bonds. *3 equiv of alkene was used. †30 mol % of CuCl and 30 mol % of Phen were used. ‡2.5 equiv of alkene was used. §5.0 equiv of alkene was used. ¶The reaction was performed at 60°C. OBn, benzyloxy; Ph, phenyl.

coupling products in good yields. Notably, aryl halides were unaffected under the reaction conditions and might be used as functional handles for further derivatization.

To demonstrate the utility of this protocol in drug discovery, we applied our dealkenylative C–N coupling strategy to the late-stage functionalization of complex bioactive molecules. An *N*-tosyl prolinamide (**25**) and a dexibuprofen amide (**26**) were decorated with chiral 2-hydroxy-4-methylcyclohexyl groups in good yields. Two known pharmaceuticals, the non-steroidal anti-inflammatory celecoxib (**27**) and the antiretroviral lamivudine (**28**), were also N-alkylated effectively. Encouraged by the exclusive alkylation of the indole nitrogen atom of tryptophan, we explored the potential for indole N-functionalization of the tryptophan

residue in the context of peptides featuring several amide N–H bonds. Both the tripeptide **29** and an endogenous opioid neurotransmitter, the endomorphin precursor **30**, underwent the site-selective C–N couplings in good yields with gentle heating. We also tested aliphatic amines, including benzylamine, dibenzylamine, and 1-phenylpiperazine, but none of them afforded desired C–N coupling products. We surmise that aliphatic amines serve as strongly donating ligands on copper, thereby influencing the coordination structures and redox potentials of the copper complexes **91** and **92**, in turn leading to inefficient regeneration of copper(I) species (see mechanistic studies and section 10.1 of the supplementary materials for details).

Scope of alkenes

Beyond (–)-isopulegol, we subjected an array of terpenes, terpenoids, and their derivatives to our deconstructive C–N coupling protocol to afford various artificial terpenoid alkaloids (Fig. 3). Most of these substrates are single enantiomers derived from natural products or are natural products themselves. (+)-Nootkatone and a eudesmane-type sesquiterpenoid underwent fragmentative C–N couplings with 3-chloroindazole and phthalimide to give the products **31a** and **31b** or **32a** and **32b**, respectively, in good yields with excellent diastereoselectivities. The bridged cyclic ketoester **33** was obtained in moderate yield and diastereoselectivity from the synthetic intermediate leading to (+)-seychellene. Although good yields could be obtained regardless of substrate, the diastereoselectivity varied widely, especially when engaging the monocyclic terpenoids (–)-limonene-1,2-diol, (+)-dihydrocarvone ethylene glycol acetal, *cis*-(+)-limonene oxide, (+)-dihydrocarveol, (–)-perillyl alcohol oxide, (–)-carveol oxide, and (–)-*O*-benzoyl isopulegol (**34** to **40**). The observed diastereoselectivities are consistent with stereoselectivity trends commonly encountered in reactions with cyclic radicals, in which the stereoselectivity of the addition is dictated by a combination of torsional and steric effects (*38*, *39*). All of the diastereoisomers [except two: the diastereoisomers of the minor regioisomers, **57a** and **57b**, from the aminodealkenylation of (–)-β-pinene] were separable through silica-column flash chromatography—a particularly relevant feature for lead compound discovery because diastereoisomers often exhibit drastically different bioactivities. Simple cyclohexyl and 4-piperidinyl radicals worked well when applied to the functionalization of melatonin (**41**) and tryptophan (**42**), respectively. We evaluated substrates bearing gem-disubstituted olefins, beyond the isopropenyl group, in the reaction and found that they provided primary alkyl-substituted amines in a facile manner (**43** to **47**). The epoxy indazole **43** was obtained in 83% yield from a (±)-α-ionone oxide derivative. A (–)-sclareol–derived alkene underwent smooth dealkenylative amination with 3-chloroindazole to afford the drimane meroterpenoid derivative **44** in 61% yield. Drimane meroterpenoids have diverse bioactivities, and our protocol provides a strategy toward N-heterocycles substituted with these terpenoid analogs (*40*). We used (*R*)-2,4-dimethylpent-4-en-1-ol, which can be accessed readily from Evans' auxiliary (~$3/g for *R*- or *S*-enantiomer) and 2-methallyl chloride, to synthesize (*R*)-3-(3-chloroindazolyl)-2-methylpropanol *O*-acetate (**45**), a precursor of muscarinic agonists, in 74% yield (*41*). Enantiopure 3-indazolyl-2-methylpropanols appear in 22 international patents and two US patents and were prepared from the corresponding (+)- and (–)-3-bromo-2-methyl-1-propanols (~$190/g from Sigma-Aldrich), which have previously been prepared from camphor sulfonic acids in five synthetic steps (*42*). Primary alkyl radicals from functionalized alkenes underwent efficient amination (**46** and **47**).

Monosubstituted alkenes were also competent substrates for our Cu-catalyzed dealkenylative C–N coupling (R = R′ = H); in contrast, these alkenes performed only moderately in Fe(II)-mediated dealkenylative radical couplings (*29*). We obtained a 73% isolated yield of the dealkenylative C–N coupling product **48** from (±)-dihydromyrcenol. Monosubstituted alkenes derived from (–)-β- citronellol and (+)-geraniol oxide *O*-acetate derivatives afforded the fragmented C–N coupling products **49a**, **49b**, and **50** in good yields. 1-Decene, which is produced from the Ziegler process or through cracking in the petroleum industry, worked well as a substrate to generate 3-chloro-1-octylindazole (**51**) in 67% yield. We found that an alkyl bromide, which could serve as a radical precursor, was tolerated under our conditions to produce *N*-(4-bromobutyl)phthalimide (**52**) in 55% yield. We also applied our deconstructive C–N coupling protocol to the modification of molecules of interest to materials science. The vinyl groups of chiral molecules used in liquid crystals were replaced by 3-chloroindazole in good yields (**53** and **54**). Another natural product, phytol, containing a trisubstituted alkene unit, underwent regioselective C(sp^3)–C(sp^2) σ bond cleavage to afford the coupling products **55a** and **55b** in good yields. Compounds **55a** and **55b** could be considered azacycle analogs of vitamin E, where the benzopyran unit is substituted by chloroindazole and phthalimide moieties, respectively.

Alkylidenecycloalkanes were also suitable substrates for the dealkenylative C–N coupling process. (–)-β-Pinene provided the synthetically valuable chiral cyclobutanes **56a** and **57a** or **56b** and **57b** as 1.7:1 and 2.9:1 mixtures of regioisomers, respectively. In contrast, (±)-sabinene afforded the cyclopropane **58** as a single regioisomer in 78% yield. Methyleneadamantane (**59**) provided the equatorial amination product **60** exclusively. The β-amino acid derivative **62** was formed smoothly from the corresponding 4-methylenepiperidine **61**. The copper-catalyzed amination also converted the trisubstituted alkene **63** into the corresponding dealkenylated C–N coupling product **64** in 88% yield.

Chiral primary amine synthesis and its application

The ability to use aminodealkenylation to install phthalimide units into chiral-pool molecules provided an opportunity to synthesize chiral primary amines (Fig. 4). For example, we prepared the primary amine **65** in 75% yield from the phytol-derived phthalimide **55b**. Chiral amino alcohols are common in nature, and such bioactive compounds often serve as auxiliaries in organic synthesis. 1,2-Aminocyclohexanol is an important motif in organic synthesis and medicinal chemistry; we prepared its unusual 5-methylated analog **66** in 79% yield upon hydrazine-mediated deprotection of the phthalimide **18**. Similarly, we synthesized the aminocyclohexandiol **67** from the opposite enantiomer of **34b** (major diastereoisomer). Although 5-amino-3-methylpentan-1-ol is a useful building block, no asymmetric synthesis has been reported previously, and its racemic form costs $965.90/g. We obtained (*R*)-**68** upon deprotection of the phthalimide **49b** derived from (–)-β-citronellol. The chiral amino acid ester **69** can be synthesized from (–)-β-pinene. The abundant functional groups in terpenoids also facilitated the synthesis of more-elaborate amino alcohols through diversification. For example, we prepared the amino alcohol **71** from *cis*-(–)-limonene oxide through aminodealkenylation, thiolation, and deprotection. We obtained the amino alcohol **73**, bearing a nootkatol skeleton, through the Luche reduction of **31b**, followed by deprotection. We synthesized the artificial indole alkaloid **75** through sequential Fischer indolization of **32b** with 4-methoxylphenylhydrazine and deprotection. (1*R*,2*R*,5*R*)-5-Amino-2-methylcyclohexanol (**77**) is a synthetic intermediate leading to a JNK inhibitor (Fig. 1D) (*24*, *43*); previous routes to its synthesis have involved 12 steps from (–)- limonene and 11 steps from the Diels-Alder reaction of isoprene and methyl acrylate. Using our deconstructive C–N coupling strategy with commercially available (–)-dihydrocarveol ($3.10/g) and phthalimide, followed by deprotection using hydrazine monohydrate, afforded **77** in two steps in 54% overall yield. Chiral 4-substituted-1-methylcyclohexenes are synthetically challenging molecules and serve as key intermediates in the synthesis of various pharmaceuticals and natural products (*44*). (*S*)-1-Methylcyclohexenyl-4-amine (**79**, $1930.50/g) has been a precursor in the synthesis of a trichodiene synthase inhibitor. The

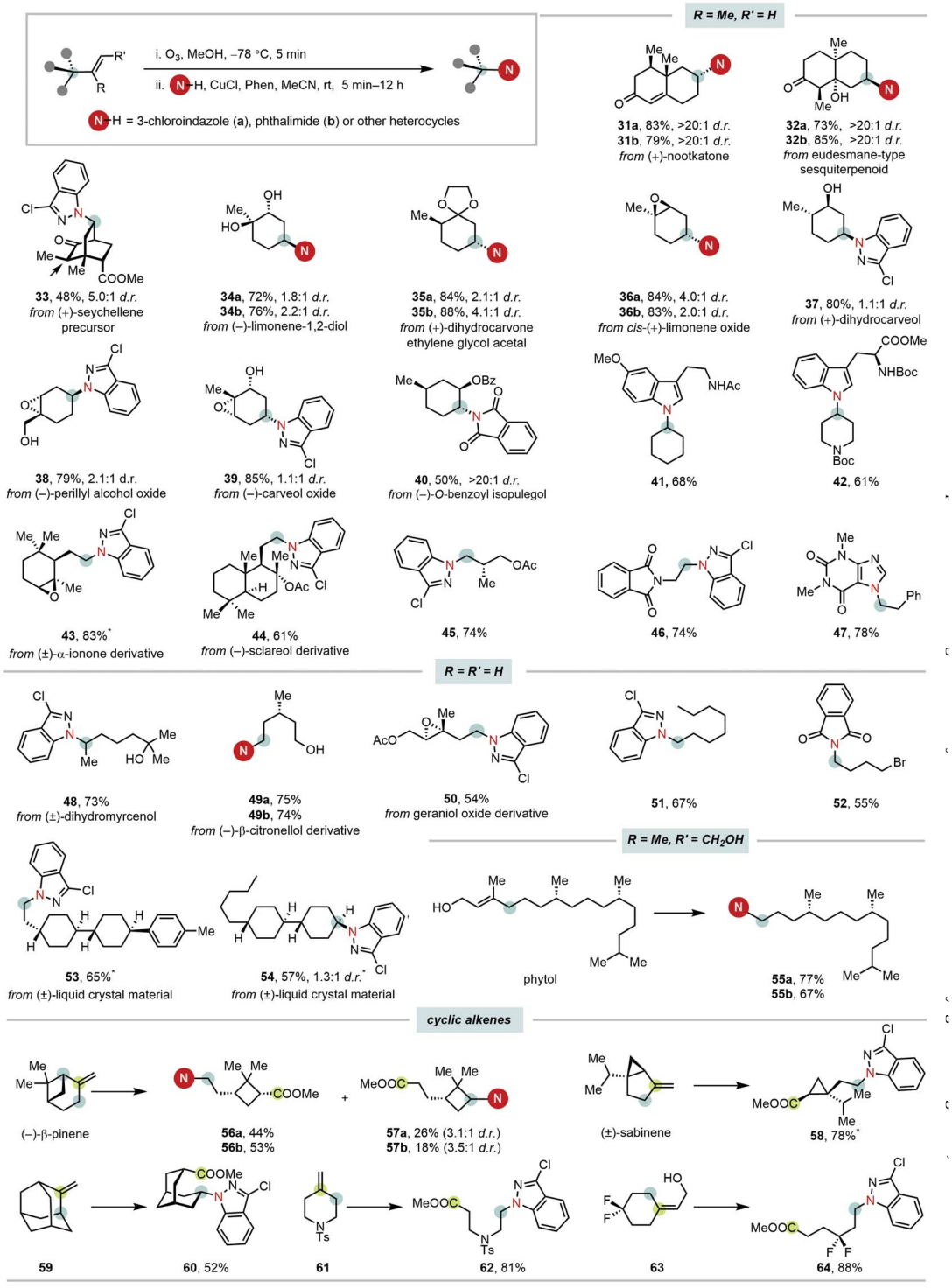


Fig. 3. Substrate scope of alkenes. All yields are isolated yields. For the detailed reaction conditions of each substrate, see the supplementary materials. Absolute configurations are indicated by wedged or dashed bonds, unless otherwise specified. The diastereoisomers were separated by flash column chromatography; major diastereoisomers are shown. *Wedged and dashed bonds indicate the relative stereochemistry, because the racemic starting material was used.

previous route to **79** involved eight steps from chiral limonenes (*45*). The enantioenriched phthalimide **78** was prepared recently by Liu and co-workers in four steps, including an alkene desymmetrization strategy using a chiral cobalt (II) complex catalyst, which was prepared in six steps from L-alaninol (*44*). Our strategy engaged commercially available *cis*-(–)-limonene oxide for dealkenylative amination, in which the C–N coupling products were produced in 83% yield with 2.1:1 diastereomeric ratio (d.r.), with the desired diastereoisomer (isolated in 56% yield), after subsequent deoxygenation, affording the chiral phthalimide **78** in 55% yield. The deprotection proceeded smoothly to afford the primary amine **79** in 85% yield (29% overall yield over three steps). Considering the capability of copper-catalyzed coupling with various nucleophiles, our strategy might provide facile access to chiral 4-substituted-1-methylcyclohexenes.

Methylation reactions

α-Methylstyrene (AMS), a by-product of the cumene process, is produced in quantities of >292,000 tons annually (*46*). Because of the relative stability of the methyl radical (BDE:

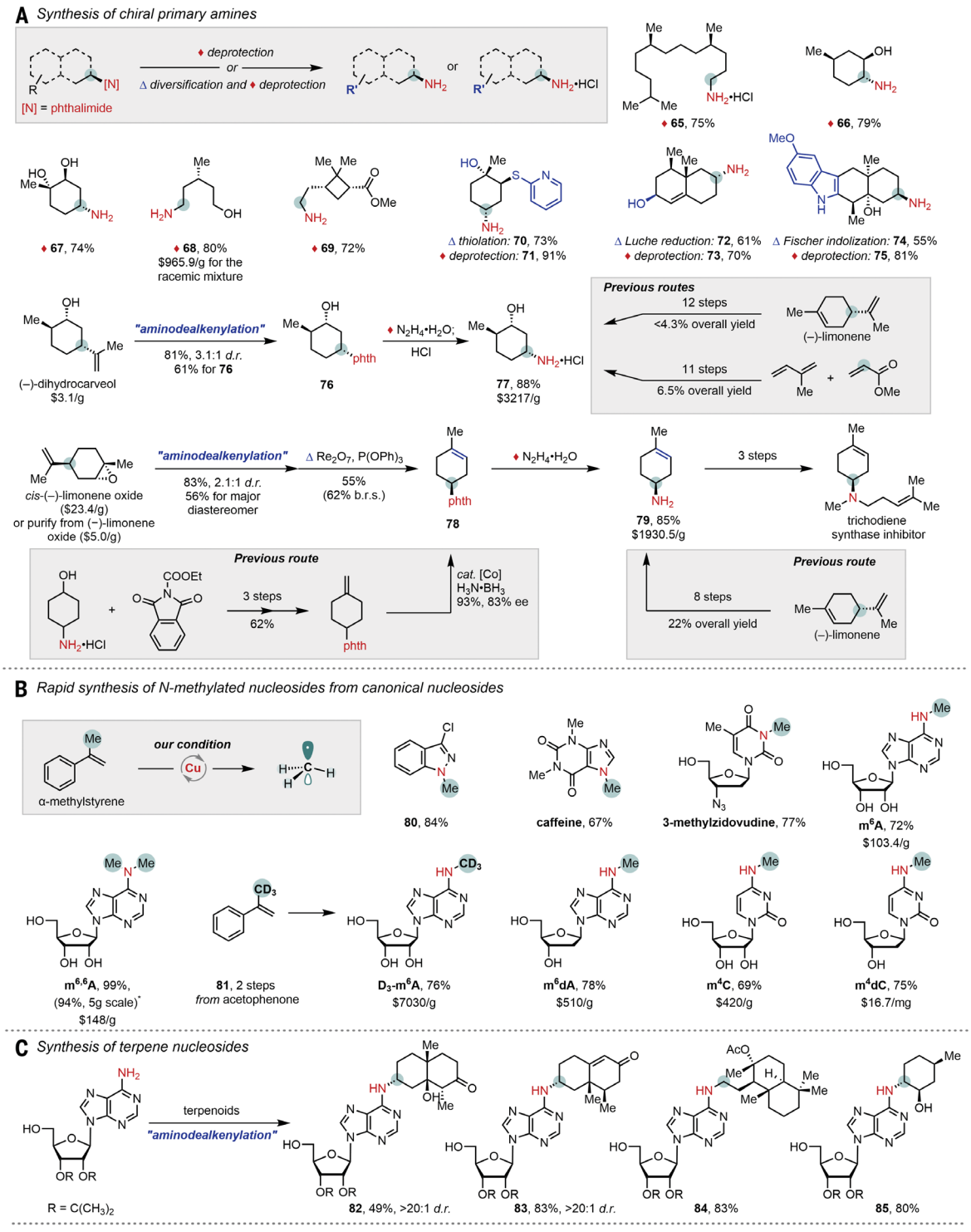


Fig. 4. Application of dealkenylative C(sp^3)–N bond coupling in bioactive compound synthesis. (**A**) Syntheses of complex chiral amines and bioactive compounds. (**B**) Methylation of nucleosides. (**C**) Terpene-nucleoside synthesis. *No flash column chromatography purification was necessary. b.r.s., base on recovered starting material; ee, enantiomeric excess.

105.0 kcal/mol for Me–H) over the phenyl radical (BDE: 112.9 kcal/mol for Ph–H) (*25*), we surmised that the Criegee ozonolysis product of AMS, upon exposure to the Cu(I) complex, should produce methyl radicals and, thereby, methylated amines (*23*). Indeed, we found that reactions with AMS readily generated 1-methyl-3-chloroindazole (**80**) and caffeine (from theophylline) under our protocol, in yields of 84 and 67%, respectively. We methylated zidovudine, an anti-HIV drug, at the N3 atom of the thymidine moiety to afford 3-methylzidovudine, without protecting the 5′-hydroxyl group. Particularly useful methylations were those of canonical nucleosides. N^6-Methyladenosine (**m^6A**) is the most abundantly modified nucleoside in eukaryotic mRNAs; it plays a crucial role in controlling gene expression in cellular, developmental, and disease processes (*47*). When using our protocol, adenosine ($0.097/g) can be methylated directly to **m^6A** ($103.40/g) in 72% isolated yield in a single step, compared with the two or three steps in previous reports. We achieved another common RNA

modification, **m^{6,6}A** ($148/g), in nearly quantitative yield when using an additional equivalent of AMS (compared with the two steps in a previous report). We also performed the reaction on a 5-g scale, providing a pure sample without the need for silica-column purification. We used the deuterium-labeled α-trideuteromethylstyrene (**81**), obtained from acetophenone and deuterium oxide in two steps, to generate isotopically labeled **D$_3$-m^{6}A** ($7030/g). We prepared other common methylated nucleosides, namely **m^{6}dA** ($510/g), **m^{4}C** ($420/g), and **m^{4}dC** ($16.70/mg), from the corresponding canonical deoxyadenosine, cytidine, and deoxycytidine, respectively, in good yields, where multiple synthetic steps or low yields were previously necessary (see the supplementary materials for previous routes).

Terpene nucleosides

Encouraged by the successful nucleoside methylation, we used aminodealkenylation for the rapid construction of terpene nucleosides (Fig. 4C). Terpene nucleoside natural products are generated through the adenosine processing of sesquiterpenoids or diterpenoids. They are found in bacterial metabolites and exhibit considerable biofunctionality and bioactivity—for example, as an antacid signaling molecule or as a component in the cell wall of *Mycobacterium tuberculosis* (*48*). Furthermore, an adenosine-processed eudesmane-type sesquiterpenoid isolated from the myxobacterium *Sorangium cellulosum* exhibited moderate antibacterial activity against a wide range of bacterial strains (*49*). We used the eudesmane-type sesquiterpenoid (**82**), (+)-nootkatone (**83**), drimane sesquiterpenoid (**84**), and the monoterpene (–)-isopulegol (**85**) to prepare terpene nucleosides in moderate to good yields. Analogs of terpene nucleosides can be applied as chemical probes of biological function, and their innate bioactivity might also provide a platform for developing lead compounds in medicinal research.

Mechanistic studies

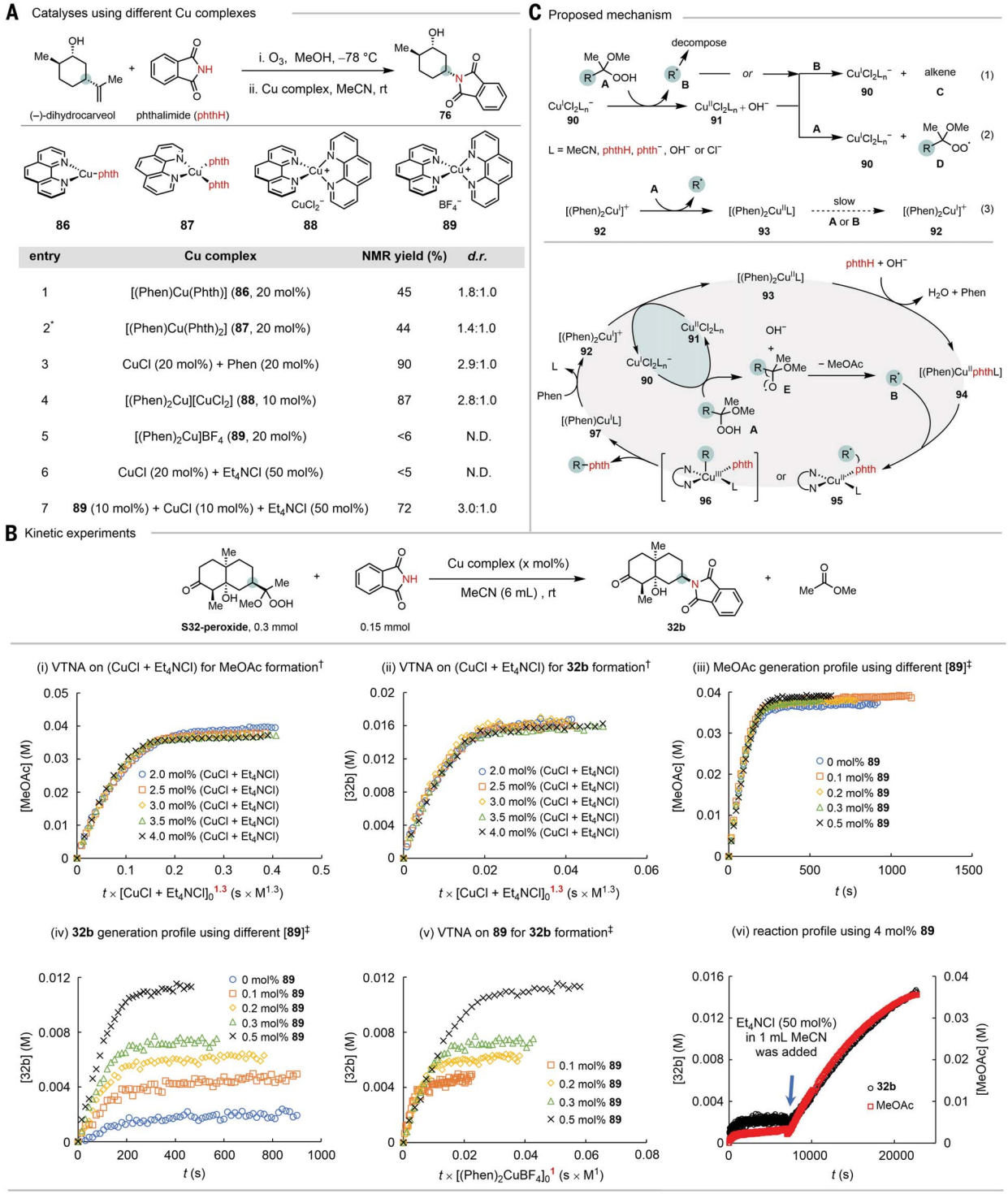
We conducted several experiments to gain insight into the reaction mechanism. Copper halides and phenanthroline are known to form tight ion pairs $[(Phen)_2Cu]^+[CuX_2]^-$ instead of neutral complexes [(Phen)CuX] in polar solvents (*50*, *51*), and our ultraviolet–visible spectroscopic measurements indicated that the ion pair complex $[(Phen)_2Cu]^+[CuCl_2]^-$ was the dominant species in MeCN (see the supplementary materials for details). Previous mechanistic studies have suggested that the neutral complex is the active species in the C–N coupling reaction, due to inefficient oxidative addition of ionic copper complexes to aryl halides (*17*, *51*). In contrast to oxidative addition, single-electron transfer (SET) between hydroperoxide and either $[(Phen)_2Cu]^+$ or $[CuCl_2]^-$ is rapid [the rate constant (k) ≈ 4×10^3 M^{-1} s^{-1}] and can trigger peroxide fragmentation (*52*, *53*).

In an attempt to directly probe the nature of the reactive copper species, we prepared a series of copper complexes (Fig. 5A) and compared their performance in the aminodealkenylation between (–)-dihydrocarveol and phthalimide. When the prototypical mixture of 1:1 CuCl and phenanthroline was replaced with a neutral phthalimido copper complex {[(Phen)Cu(phth)] (**86**) or [(Phen)Cu(phth)$_2$] (**87**)}, which had been proposed previously as a reactive intermediate (*15*, *17*), both the product yields (45 and 44%, respectively) and diastereoisomeric ratios (1.8:1.0 and 1.4:1.0, respectively) deteriorated substantially when compared with those of the standard reaction (90% nuclear magnetic resonance yield, 2.9:1.0 d.r.) (entries 1 to 3). Using the ion pair $[(Phen)_2Cu]^+$ $[CuCl_2]^-$ (**88**) directly afforded a reaction similar to the prototypical aminodealkenylation (87% yield, 2.8:1.0 d.r.), indicating that the ion pair might have been the active catalytic species. Reactions using either $[(Phen)_2Cu]^+$ (**89**) or $[CuCl_2]^-$ (CuCl + Et_4NCl) were extremely sluggish (entries 5 and 6), whereas the combination of independently prepared $[(Phen)_2Cu]^+$ and $[CuCl_2]^-$ afforded a product similar to that from the parent reaction (entries 4 and 7), but with a slightly diminished yield.

To gain further insight into the mechanism of the copper-catalyzed amination, we conducted kinetic studies through variable time normalization analysis (VTNA) (*54*). Using ReactIR, we collected the concentration–time profiles for both the C–C scission product [methyl acetate (MeOAc)] and the C–N coupling product (**32b**). The formation of MeOAc displayed 0.3-, zero-, and 1.3-order dependences with respect to **S32-peroxide**, phthalimide, and the catalyst (CuCl + Phen), respectively, while the coupling product **32b** displayed 0.3-, 0.3-, and 1.3-order dependences (figs. S22 to S27). The zero– and pseudo-zero–order dependences on both peroxide and the amine imply that there existed an off-cycle catalyst resting state (*55*, *56*). Because both $[(Phen)_2Cu]^+$ and $[CuCl_2]^-$ are readily oxidized by peroxide, we surmise that the resting states of the catalyst were the corresponding copper(II) species **91** and **93** (Fig. 5C, reactions 1 to 3). The kinetic order of 1.3 with respect to catalyst indicates a scenario in which two catalytic species were operating on-cycle (Fig. 5C).

To tease out the roles of the cation and anion complexes in the catalytic cycle, we investigated the kinetics of $[(Phen)_2Cu]^+$ (**89**) and $[CuCl_2]^-$ (1:1 mixture of CuCl + Et_4NCl) separately. Both the C–C scission and C–N coupling events exhibited a kinetic order of 1.3 in $[CuCl_2]^-$ in the presence of 2 mol % **89** (Fig. 5B, i and ii, and figs. S32 and S33) and a kinetic order of 2 in the presence of 0.5 mol % **89** (figs. S34 and S35). Moreover, although $[CuCl_2]^-$ alone could catalyze the generation of MeOAc (Fig. 5B, iii), the formation of **32b** was sluggish in the absence of $[(Phen)_2Cu]^+$ (Fig. 5B, iv, and Fig. 5A, entry 6). In sum, the C–C scission and C–N coupling steps displayed zero- and first-order dependences on $[(Phen)_2Cu]^+$ at low concentrations (0.1 to 0.5 mmol %; Fig. 5B, iii and v) and saturation kinetics at higher concentrations (0.5 to 2.0 mol %), respectively, each in the presence of 2.0 mol % $[CuCl_2]^-$ (figs. S28 to S31). The amount of **89** did not affect C–C scission, generating MeOAc even in its absence (Fig. 5B, iii), while both the production rate and yield of **32b** were dependent on the concentration of **89** (Fig. 5B, iv). Complex **89** alone, however, could not generate appreciable amounts of MeOAc and **32b**, both of which were generated steadily upon addition of Et_4NCl (Fig. 5B, vi). We surmise that $[(Phen)_2Cu]^{2+}$ (**93**) was a stable cation whose reduction back to Cu(I) species was sluggish. Therefore, once $[(Phen)_2Cu]^+$ was oxidized, the reaction stalled (Fig. 5C, reaction 3). Nevertheless, the addition of chloride replenished some $Cu(II)Cl_mL_n$ species that could oxidize the peroxide **A** and the radical **B** and return to the Cu(I) species (Fig. 5C, reactions 1 and 2) (*57*). The disparate reaction kinetics of $[CuCl_2]^-$ and $[(Phen)_2Cu]^+$ indicate that $[CuCl_2]^-$ was involved in hydroperoxide decomposition, whereas $[(Phen)_2Cu]^+$ participated in the C–N coupling. Because C–C scission occurred before C–N coupling, $[CuCl_2]^-$ influenced both the C–C scission and C–N coupling processes (figs. S34 and S35).

Taken together, we propose the involvement of $[CuCl_2]^-$–$[(Phen)_2Cu]^+$ cooperative catalysis (Fig. 5C). SET between complex **90** and the peroxide **A** affords the alkoxyl radical **E**, which undergoes β-scission to generate the alkyl radical **B**. The cationic Cu(I) complex **92** is oxidized by peroxide **A** to afford a copper(II) complex **93**. Deprotonated phthalimide associates with complex **93**, followed by dissociation of Phen, to afford the phthalimido copper(II) complex **94**. Complex **94** traps the alkyl radical **B** to afford the C–N coupling product and the copper(I) complex **97** through either an outer sphere (**95**) or inner sphere (**96**) pathway (*58*). Ligand exchange on the Cu(I) complex **97** affords **92**. Subsequent electron transfer between **91** and **92** regenerates **90** and **93**, completing a catalytic cycle. Alternatively, electron transfer between **97** and **91** is also possible to regenerate **90**. Deprotonated phthalimide associates with the resulting Cu(II) complex from **97** to regenerate **94**, completing the catalytic cycle. Given the straightforward nature of the modular introduction of nitrogen moieties through the unusual disconnection of alkene skeletons, we anticipate that our copper-catalyzed dealkenylative C–N coupling will find applications in the preparation of complex bioactive three-dimensional molecules and optically



entry	Cu complex	NMR yield (%)	*d.r.*
1	[(Phen)Cu(Phth)] (**86**, 20 mol%)	45	1.8:1.0
2*	[(Phen)Cu(Phth)$_2$] (**87**, 20 mol%)	44	1.4:1.0
3	CuCl (20 mol%) + Phen (20 mol%)	90	2.9:1.0
4	[(Phen)$_2$Cu][CuCl$_2$] (**88**, 10 mol%)	87	2.8:1.0
5	[(Phen)$_2$Cu]BF$_4$ (**89**, 20 mol%)	<6	N.D.
6	CuCl (20 mol%) + Et$_4$NCl (50 mol%)	<5	N.D.
7	**89** (10 mol%) + CuCl (10 mol%) + Et$_4$NCl (50 mol%)	72	3.0:1.0

Fig. 5. Mechanistic studies. (**A**) Comparison of different copper complexes. (**B**) Kinetic studies. (**C**) Proposed mechanism. *MeCN/MeOH (9:1 v/v) was used because complex **87** does not dissolve in MeCN. †2 mol % of **89** was used. ‡2 mol % of (CuCl + Et$_4$NCl) was used. N.D., not determined.

active amines of great interest in organic synthesis and medicinal chemistry.

REFERENCES AND NOTES

1. E. Vitaku, D. T. Smith, J. T. Njardarson, *J. Med. Chem.* **57**, 10257–10274 (2014).
2. D. G. Brown, J. Boström, *J. Med. Chem.* **59**, 4443–4458 (2016).
3. A. Trowbridge, S. M. Walton, M. J. Gaunt, *Chem. Rev.* **120**, 2613–2692 (2020).
4. F. Chen, T. Wang, N. Jiao, *Chem. Rev.* **114**, 8613–8661 (2014).
5. Y. Xia, G. Dong, *Nat. Rev. Chem.* **4**, 600–614 (2020).
6. J. B. Roque, Y. Kuroda, L. T. Göttemann, R. Sarpong, *Nature* **564**, 244–248 (2018).
7. J. Tinge *et al*., in *Ullmann's Encyclopedia of Industrial Chemistry* (Wiley, 2018), pp. 1–31.
8. J. Liu *et al*., *Nat. Chem.* **11**, 71–77 (2019).
9. J. Liu *et al*., *Science* **367**, 281–285 (2020).
10. J. Liu *et al*., *Research* **2020**, 7947029 (2020).
11. W. Zhao, R. P. Wurz, J. C. Peters, G. C. Fu, *J. Am. Chem. Soc.* **139**, 12153–12156 (2017).
12. R. Mao, A. Frey, J. Balon, X. Hu, *Nat. Catal.* **1**, 120–126 (2018).
13. Y. Liang, X. Zhang, D. W. C. MacMillan, *Nature* **559**, 83–88 (2018).

14. X. Hu *et al*., *ACS Catal.* **10**, 8402–8408 (2020).
15. X.-Y. Lv, R. Abrams, R. Martin, *Angew. Chem. Int. Ed.* **62**, e202217386 (2023).
16. G. P. Chiusoli, F. Minisci, *Gazz. Chim. Ital.* **88**, 261–270 (1958).
17. B. L. Tran, B. Li, M. Driess, J. F. Hartwig, *J. Am. Chem. Soc.* **136**, 2555–2563 (2014).
18. J.-J. Guo *et al*., *Angew. Chem. Int. Ed.* **55**, 15319–15322 (2016).
19. S. Sakurai, T. Kato, R. Sakamoto, K. Maruoka, *Tetrahedron* **75**, 172–179 (2019).
20. J. Bao, H. Tian, P. Yang, J. Deng, J. Gui, *Eur. J. Org. Chem.* **2020**, 339–347 (2020).
21. P. Ertl, T. Schuhmann, *J. Nat. Prod.* **82**, 1258–1263 (2019).
22. T. C. Nugent, Ed., *Chiral Amine Synthesis: Methods, Developments and Applications* (Wiley-VCH, 2010).
23. F. Ospina *et al*., *Angew. Chem. Int. Ed.* **61**, e202213056 (2022).
24. H.-W. Man *et al*., Solid forms of 2-(*tert*-butylamino)-4-((1*R*,3*R*,4*R*)-3-hydroxy-4-methylcyclohexylamino)-pyrimidine-5-carboxamide, compositions thereof and methods of their use, United States Patent US 9365524 B2, 14 June 2016.
25. Y.-R. Luo, *Comprehensive Handbook of Chemical Bond Energies* (CRC Press, 2007).
26. T. J. Fisher, P. H. Dussault, *Tetrahedron* **73**, 4233–4258 (2017).
27. E. I. Epelle *et al*., *Chem. Eng. J.* **454**, 140188 (2023).
28. R. Criegee, G. Wenner, *Justus Liebigs Ann. Chem.* **564**, 9–15 (1949).
29. A. J. Smaligo *et al*., *Science* **364**, 681–685 (2019).
30. J. H. Dworkin, B. D. Dehnert, O. Kwon, *Trends Chem.* **5**, 174–200 (2023).
31. B. D. Dehnert, J. H. Dworkin, O. Kwon, *Synthesis* (2023).
32. R. Trammell, K. Rajabimoghadam, I. Garcia-Bosch, *Chem. Rev.* **119**, 2954–3031 (2019).
33. A. C. Bissember, R. J. Lundgren, S. E. Creutz, J. C. Peters, G. C. Fu, *Angew. Chem. Int. Ed.* **52**, 5129–5133 (2013).
34. C. Chen, J. C. Peters, G. C. Fu, *Nature* **596**, 250–256 (2021).
35. B. Górski, A.-L. Barthelemy, J. J. Douglas, F. Juliá, D. Leonori, *Nat. Catal.* **4**, 623–630 (2021).
36. N. W. Dow, A. Cabré, D. W. C. MacMillan, *Chem* **7**, 1827–1842 (2021).
37. T. V. Sravanthi, S. L. Manju, *Eur. J. Pharm. Sci.* **91**, 1–10 (2016).
38. B. Giese, *Angew. Chem. Int. Ed.* **28**, 969–980 (1989).
39. G. Bar, A. F. Parsons, *Chem. Soc. Rev.* **32**, 251–263 (2003).
40. D. D. Dixon, J. W. Lockner, Q. Zhou, P. S. Baran, *J. Am. Chem. Soc.* **134**, 8432–8435 (2012).
41. E. S. Burstein *et al*., Preparation of indazole compounds as muscarinic agonists, International Patent WO 2014/152144 Al, 25 September 2014.
42. S. Mekala, R. C. Hahn, *J. Org. Chem.* **80**, 1610–1617 (2015).
43. T. J. Connolly, H.-W. Man, P. Nagarajan, C. Rajendiran, J. Venkateswarlu, Methods of synthesis of (1*R*,2*R*,5*R*)-5-amino-2-methylcyclohexanol hydrochloride and intermediates useful therein, International Patent WO 2017/019487 A1, 2 February 2017.
44. X. Liu, X. Rong, S. Liu, Y. Lan, Q. Liu, *J. Am. Chem. Soc.* **143**, 20633–20639 (2021).
45. D. E. Cane, G. Yang, R. M. Coates, H. J. Pyun, T. M. Hohn, *J. Org. Chem.* **57**, 3454–3462 (1992).
46. The annual production volume of α-methylstyrene (AMS) was obtained at https://www.globenewswire.com/news-release/2019/05/27/1850397/0/en/542-9-Mn-Alpha-Methylstyrene-Market-Global-Forecast-to-2024.html. Please refer to the supplementary materials for its price as well as the price of each chemical in this study.
47. S. Zaccara, R. J. Ries, S. R. Jaffrey, *Nat. Rev. Mol. Cell Biol.* **20**, 608–624 (2019).
48. M. Holzheimer, J. Buter, A. J. Minnaard, *Chem. Rev.* **121**, 9554–9643 (2021).
49. J.-W. Ahn, K. H. Jang, S.-C. Chung, K.-B. Oh, J. Shin, *Org. Lett.* **10**, 1167–1169 (2008).
50. A. J. Pallenberg, K. S. Koenig, D. M. Barnhart, *Inorg. Chem.* **34**, 2833–2840 (1995).
51. J. W. Tye, Z. Weng, A. M. Johns, C. D. Incarvito, J. F. Hartwig, *J. Am. Chem. Soc.* **130**, 9971–9983 (2008).
52. M. Masarwa *et al*., *J. Am. Chem. Soc.* **110**, 4293–4297 (1988).
53. K. V. Ponganis, M. A. De Araujo, H. L. Hodges, *Inorg. Chem.* **19**, 2704–2709 (1980).
54. J. Burés, *Angew. Chem. Int. Ed.* **55**, 2028–2031 (2016).
55. D. G. Blackmond, *J. Am. Chem. Soc.* **137**, 10852–10866 (2015).
56. C. Alamillo-Ferrer, G. Hutchinson, J. Burés, *Nat. Rev. Chem.* **7**, 26–34 (2023).
57. J. W. Moffett, R. G. Zika, *Environ. Sci. Technol.* **21**, 804–810 (1987).
58. H. Lee *et al*., *J. Am. Chem. Soc.* **144**, 4114–4123 (2022).

ACKNOWLEDGMENTS

We thank the UCLA Molecular Instrumentation Center for performing nuclear magnetic resonance spectroscopy and mass spectrometry. We thank Mettler Toledo for providing ReactIR and S. Paul for technical instruction. **Funding:** Financial support for this study was provided by the NIH (grant R01GM141327). **Author contributions:** Z.H. and O.K. conceived of the study, designed the experiments, and analyzed the data. Z.H., J.A.M., M.S., and J.W. performed the experiments. Z.H. and O.K. wrote the manuscript. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** Experimental procedures, optimization data, and characterization data are available in the supplementary materials.

License information: Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adi4758
Materials and Methods
Supplementary Text
Figs. S1 to S47
Tables S1 to S7
Characterization Data
References (*59*–*129*)

Submitted 27 April 2023; accepted 25 July 2023
10.1126/science.adi4758

CHEMICAL TECHNOLOGY

Tandem propane dehydrogenation and surface oxidation catalysts for selective propylene synthesis

Wei Wang[1,2,3], Sai Chen[1,3], Chunlei Pei[1,3], Ran Luo[1,2,3], Jiachen Sun[1,3], Hongbo Song[1,3], Guodong Sun[1,3], Xianhui Wang[1,3], Zhi-Jian Zhao[1,3], Jinlong Gong[1,2,3]*

Direct propane dehydrogenation (PDH) to propylene is a desirable commercial reaction but is highly endothermic and severely limited by thermodynamic equilibrium. Routes that oxidatively remove hydrogen as water have safety and cost challenges. We coupled chemical looping–selective hydrogen (H_2) combustion and PDH with multifunctional ferric vanadate–vanadium oxide ($FeVO_4$-VO_x) redox catalysts. Well-dispersed VO_x supported on aluminum oxide (Al_2O_3) provides dehydrogenation sites, and adjacent nanoscale $FeVO_4$ acts as an oxygen carrier for subsequent H_2 combustion. We achieved an integral performance of 81.3% propylene selectivity at 42.7% propane conversion at 550°C for 200 chemical looping cycles for the reoxidization of $FeVO_4$. Based on catalytic experiments, spectroscopic characterization, and theory calculations, we propose a hydrogen spillover–mediated coupling mechanism. The hydrogen species generated at the VO_x sites migrated to adjacent $FeVO_4$ for combustion, which shifted PDH toward propylene. This mechanism is favored by the proximity between the dehydrogenation and combustion sites.

Propylene, a primary building block in the chemical industry, is mostly produced through the stream cracking of naphtha (*1*), an energy-intensive process that relies on petroleum feedstocks. Propane from shale gas is converted commercially to propylene through propane dehydrogenation (PDH) (*2*–*4*), but the reaction is highly endothermic and equilibrium-limited and requires frequent catalyst regeneration after coke deposition. An alternative, thermocatalytic oxidative dehydrogenation (ODH) of propane (*5*–*8*) must overcome the overoxidation of propylene to CO_2 in order to become a viable commercial process. Based on Le Chatelier's principle, selective removal of H_2 from products would shift the equilibrium toward propylene. Compared with hydrogen permselective membrane reactors and metal hydrides (*9*, *10*), selective hydrogen combustion (SHC) has the additional advantage of simultaneous thermal release and can be seamlessly integrated with existing industrial PDH processes (*11*, *12*). However, the O_2-cofeed catalytic systems suffer from inherent safety hazards and require energy-intensive air separation (fig. S1).

The lattice oxygen of reducible metal oxides can act as solid oxygen carriers to achieve SHC under anaerobic conditions (*13*–*15*). The feasibility of using oxygen carriers, such as Bi_2O_3 (*16*), In_2O_3 (*17*), V_2O_5 (*18*), Fe_2O_3 (*19*), MoO_3 (*20*), and doped CeO_2 (*21*), has been extensively validated. However, several limitations, including poor redox stability, low oxygen storage capacity (OSC), and limited hydrogen combustion selectivity, have precluded potential commercial use. Additionally, achieving precise control of communication behavior between dehydrogenation sites and hydrogen combustion sites at the molecular level to mediate tandem PDH-SHC reactions remains a challenge (*1*, *12*).

We propose an intensified process for coupling PDH and chemical looping–selective hydrogen combustion (CL-SHC) under anaerobic conditions in a tandem fashion with the multifunctional $FeVO_4$-VO_x redox catalysts (Fig. 1A). We used a known PDH catalyst, VO_x with 3 wt % metal loading on a γ-Al_2O_3 support (denoted 3V/Al), with a synthesized $FeVO_4$ oxygen carrier that has a high OSC and high stability for SHC loaded in one reactor (figs. S2 and S3). The $FeVO_4$ oxygen carriers donated lattice oxygen to convert hydrogen into water. The reduced oxygen carriers were then exposed to air in a separate step to replenish the lattice oxygen, which is known as the chemical-looping (or redox) process. In this way, we avoided the direct contact of gaseous O_2 with the flammable mixture and the need for cryogenic air separation.

Tandem catalyst systems

We investigated four different tandem modes with site proximities ranging from the millimeter scale to the nanoscale (denoted as M1 to M4) by controlling the spatial distribution of $FeVO_4$ oxygen carriers and 3V/Al PDH catalysts (Fig. 1B and fig. S4). All materials were pressed into pellets and sieved to prepare granules with sizes of 400 to 800 μm before loading into the reactor (see details in the supplemental materials). In M1, we used a dual-bed configuration in which $FeVO_4$ was packed below 3V/Al, separated by a layer of inert quartz sand. In M2, we mixed the granules with $FeVO_4$ and 3V/Al well. In M3, we used mixtures of $FeVO_4$ and 3V/Al fine powders (50 to 75 μm), which were then pressed, crushed, and sieved to granules ranging in size from 400 to 800 μm. In M4, we mortar-ground mixtures of $FeVO_4$ and 3V/Al powders to achieve nanoscale intimacy, where material transferred between particles so that both catalyst and oxygen carrier were on the surface of the same particle (denoted *x*FeV-3V/Al, where *x* represents the mass percentage of $FeVO_4$ in the tandem catalysts) (figs. S5 and S6 and table S1).

Structural analysis of 30FeV-3V/Al tandem catalysts with nanoscale intimacy

We confirmed the feasibility of the structural hypothesis that 30FeV-3V/Al tandem catalysts (M4) achieved nanoscale intimacy with complementary characterization techniques. Powder x-ray diffraction patterns showed that 30FeV-3V/Al was composed of triclinic $FeVO_4$ phase (PDF#25-0418) and γ-Al_2O_3 (Fig. 1C). The absence of vanadium oxide peaks implied the high dispersion of VO_x species (fig. S7).

The in situ Raman spectrum of dehydrated 30FeV-3V/Al contained a set of characteristic peak bands between 200 and 1000 cm^{-1} that we attributed to bulk $FeVO_4$ (*22*) and a characteristic band at about 1020 cm^{-1} that we attributed to V=O bonds in the highly dispersed VO_x species on the surface of the γ-Al_2O_3 support (Fig. 1D) (*23*). High-resolution transmission electron microscopy images and energy-dispersive x-ray spectroscopy mappings also showed independent $FeVO_4$ nanoparticles (NPs) and highly dispersed VO_x on Al_2O_3 (3V/Al) with nanoscale intimacy (figs. S8 and S9).

Electron energy-loss spectroscopy (EELS) of the V $L_{2,3}$ and Fe $L_{2,3}$ edges crossing the interface of $FeVO_4$ NPs and 3V/Al was used to identify the spatial distribution of the matrix (Fig. 1E), where three different zones were selected (Fig. 1F): (i) $FeVO_4$ NPs, (ii) 3V/Al adjacent to $FeVO_4$ NPs, and (iii) 3V/Al away from $FeVO_4$ NPs. The variations of the intensity of the Fe $L_{2,3}$ edges from zone 1 to zones 2 and 3 confirmed that the Fe atoms were exclusively concentrated in the $FeVO_4$ NPs (Fig. 1G), as shown in the EELS elemental mappings of Fe and V (Fig. 1F).

The V $L_{2,3}$ edges showed two adsorption jumps at $E(L_3)$ = 518.6 eV and $E(L_2)$ = 525.4 eV on the $FeVO_4$ NPs (zone 1), illustrating the dominance of V^{5+} in $FeVO_4$. The V L_3 edge at the adjacent 3V/Al (zones 2 and 3) shifted toward lower energy compared with that at the $FeVO_4$ NPs (zone 1), indicating that VO_x species were in a mixed V^{x+} (x = 3 to 5) valence state (Fig. 1H) (*24*), which was also evidenced

[1]Key Laboratory for Green Chemical Technology of Ministry of Education, School of Chemical Engineering and Technology, Tianjin University, Tianjin 300072, China. [2]Joint School of National University of Singapore and Tianjin University, International Campus of Tianjin University, Binhai New City, Fuzhou 350207, China. [3]Collaborative Innovation Center for Chemical Science and Engineering (Tianjin), Tianjin 300072, China.
*Corresponding author. Email: jlgong@tju.edu.cn

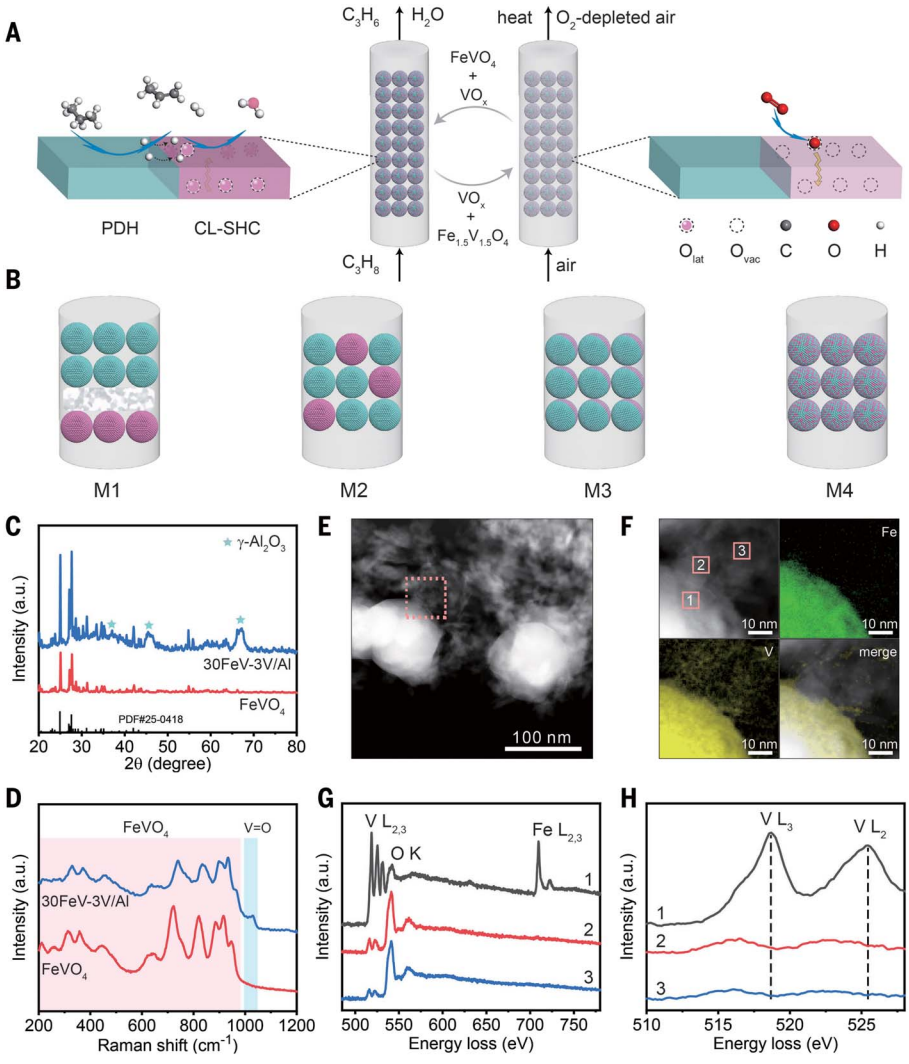


Fig. 1. Characterizations of 30FeV-3V/Al tandem catalysts. (**A**) Schematic representation of the PDH+CL-SHC scheme under anaerobic conditions. lat, lattice; vac, vacancy. (**B**) Schematic of the tandem catalysts in different modes: M1, dual-bed configuration with $FeVO_4$ packed below 3V/Al; M2, stacking of granules with $FeVO_4$ and 3V/Al; M3, powder mixtures of $FeVO_4$ and 3V/Al; and M4, mortar mixtures of 3V/Al and $FeVO_4$. (**C** and **D**) XRD patterns (C) and in situ Raman spectra (D) of $FeVO_4$ and 30FeV-3V/Al. a.u., arbitrary units. (**E** and **F**) HAADF-STEM image (E) and EELS-mapping (F) of 30FeV-3V/Al, which comes from the selected square area in (E). (**G** and **H**) EELS line-scanning of the V $L_{2,3}$ and Fe $L_{2,3}$ edges.

by the deconvolution of V 2p x-ray photoelectron spectroscopy results (figs. S10 and S11). These characterizations support a structural model in which $FeVO_4$ NPs are surrounded by the 3V/Al matrix and the NPs and matrix are situated in close proximity at the nanoscale.

Intimacy-dependent propylene selectivity and H combustion selectivity

We demonstrated the application of $FeVO_4$-VO_x tandem catalysts in a continuous PDH+CL-SHC scheme. The $FeVO_4$ oxygen carrier alone tended to combust propane, but this reaction competed with PDH in the presence of 3V/Al catalysts (Fig. 2A and fig. S12). We observed strong correlations between the proximity of $FeVO_4$ oxygen carriers and 3V/Al catalysts and the selectivity of propylene. Integrated catalysts with a dual-bed configuration (M1), in which $FeVO_4$ was packed below the 3V/Al bed, and granular mixture of $FeVO_4$ and 3V/Al (M2) exhibited rather high CO_x selectivity (34.5%) and poor propylene selectivity (56.2%) at 5 min on stream (fig. S13). Propylene selectivity was improved to 75.7% by changing the $FeVO_4$ and 3V/Al into a closer proximity in M3 (Fig. 2A).

We found that 30FeV-3V/Al (M4), which exhibited nanoscale intimacy, attained the highest propylene selectivity (86.3%) with negligible CO_x production (3.6%) and high carbon balance (97.3%) at 46.5% propane conversion (Fig. 2C and fig. S13) at 5 min on stream, which outperformed the 3V/Al (fig. S14), $FeVO_4$, and state-of-the-art ODH, PDH+SHC, PDH, CL-ODH, PDH+CO_2, and PDH+CL-SHC catalysts (*6*, *12*, *16*, *25*–*40*). On-line mass spectrometry (MS) results indicated that the lattice oxygens of $FeVO_4$ that were involved in selective H combustion were sustained for at least 10 min during the course of PDH (fig. S15). The integral conversion of propane and propylene selectivity of the M4 catalyst with the consideration of coke deposits during the course of the 10-min reaction were determined to be 42.7 and 81.3%, respectively (fig. S16). Even at 40% propane concentration, the M4 catalysts achieved 93.6% propylene selectivity at 37.6% propane conversion (fig. S17) with a space-time yield of 14.6 mmol C_3H_6 g^{-1}_{cat} $hour^{-1}$, which offered a potential advantage of the PDH+CL-SHC scheme over conventional PDH-SHC with O_2 cofeeding catalytic systems (fig. S18).

We semiquantitatively estimated the distance between 3V/Al and $FeVO_4$ (denoted $d_{FeVO4\text{-}3V/Al}$) from the millimeter-scale to the nanoscale by a Monte Carlo method (fig. S19 and table S2) (*41*, *42*). The dependence of propylene selectivity and H combustion selectivity F_{tandem} [denoted $(1 - H_2/C_3H_6) \times 100\%$] on $d_{FeVO4\text{-}3V/Al}$ indicated the communication between tandem catalysts (Fig. 2B). This result is consistent with the "the closer, the better" notion of intimacy put forward by Weisz in 1962 (*43*).

We further demonstrated the stability and recyclability of the 30FeV-3V/Al tandem redox catalysts with 200 redox cycles at 550°C by switching propane and air as the reducing and oxidating gases, respectively (Fig. 2D). Real-time thermogravimetric analysis (TGA) showed that the oxygen release and weight loss in each cycle were maintained at values between 3.0 and 3.5% in a test lasting >300 hours (Fig. 2E), indicating the highly facile and reversible oxygen release and storage of 30FeV-3V/Al tandem catalysts with minor changes of diffraction peaks in the x-ray diffraction (XRD) patterns (figs. S20 to S22).

In addition to intimacy-dependent propylene selectivity and H combustion selectivity, the H abstraction from propane on dehydrogenation sites played a prerequisite role for this H combustion. 3V/Al contains low loading of V (3 wt %), which can be easily reduced (fig. S23), and functions as dehydrogenation sites. Temperature-programmed surface reaction of propane indicated the sluggish activation of propane on 3V/Al at the low temperature of 450°C (fig. S24), which leads to the absence of H abstraction and the subsequent site proximity independence (fig. S25). Alkali metals are selective titrants of acid-base pairs on oxide surfaces (*44*, *45*), and we used sodium to poison the acid-base pairs of the 3V/Al surface. The absence of improvements in propylene selectivity also provided compelling evidence for the

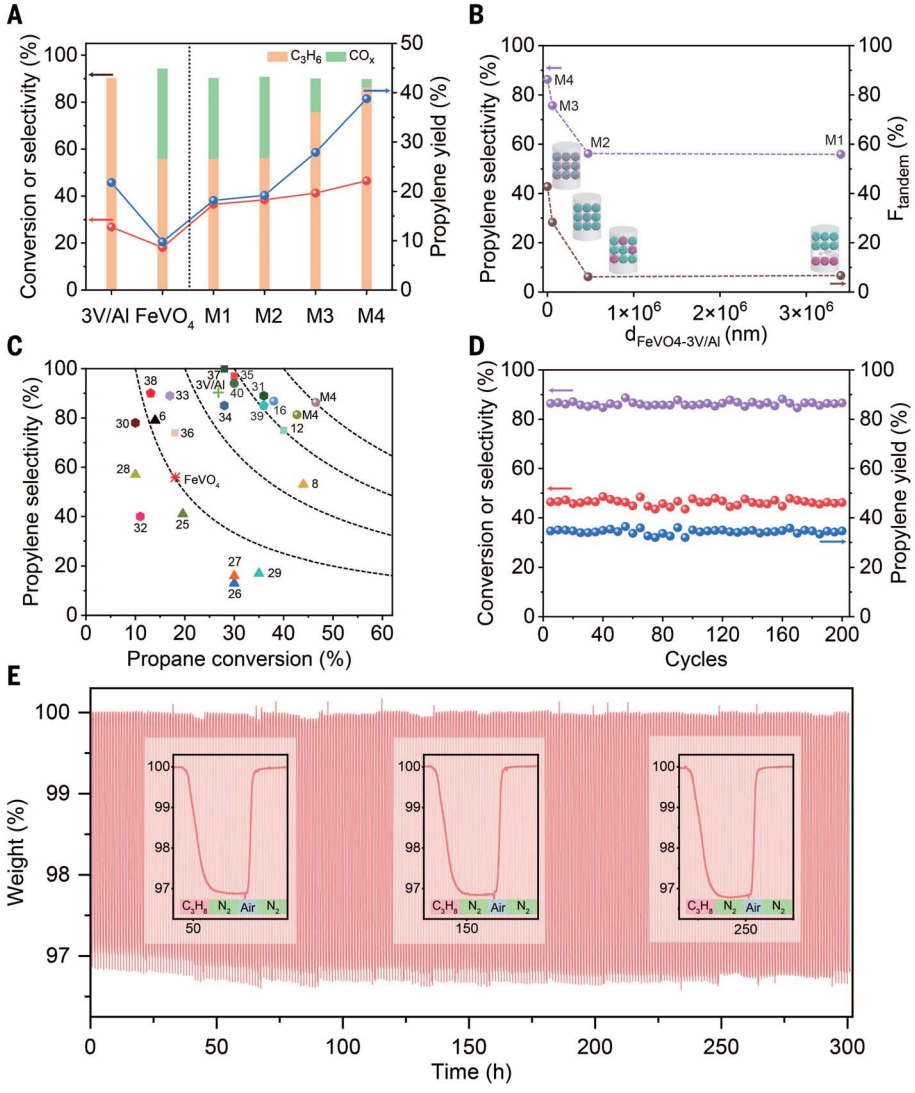


Fig. 2. Catalytic performance of tandem catalysts. (**A**) Catalytic performance of tandem catalysts with different modes at 5 min on stream [550°C, gas hourly space velocity (GHSV) = 2500 hour^{-1}, 10 vol % C_3H_8/N_2]. (**B**) Propylene selectivity and F_{tandem} as a function of $d_{FeVO4\text{-}3V/Al}$. (**C**) Comparison of 30FeV-3V/Al with best-performing reported ODH, PDH+SHC, PDH, CL-ODH, PDH+CO_2, and PDH+CL-SHC catalysts, shown as triangles, squares, pentagons, hexagons, circles, and spheres, respectively (see table S3). The brown and green spheres of M4 indicate instantaneous performance at 5 min and integral performance during the course of 10-min PDH, respectively. Numbers indicate corresponding references, as also shown in table S3. (**D**) Data from 200 redox cycle tests on 30FeV-3V/Al catalysts at 550°C. Conditions are as follows: dehydrogenation step, GHSV = 2500 hour^{-1}, 10 vol % C_3H_8/N_2 = 20 ml min^{-1} for 5 min; inert purge, N_2 = 50 ml min^{-1} for 5 min; and reoxidation step, 20 vol % O_2/N_2 = 20 ml min^{-1} for 5 min. (**E**) Real-time TGA data of 30FeV-3V/Al catalysts over >300 hours at 550°C.

tandem process of PDH with CL-SHC, in which 3V/Al played the role of catalyst to abstract hydrogen and $FeVO_4$ acted as the oxygen carrier to combust hydrogen (fig. S26). These results indicate that the intimacy-dependent propylene selectivity is closely related to the abstraction of hydrogen from propane over the PDH catalyst (3V/Al). We speculate that a surface diffusion mechanism of hydrogen species is responsible for the coupling of $FeVO_4$ and 3V/Al catalysts to complete the tandem catalytic cycle.

Mechanistic studies of the hydrogen spillover–coupling mechanism of PDH with CL-SHC

We performed H_2 temperature-programmed reduction (H_2-TPR) to investigate the redox properties of tandem catalysts and validate the spillover of hydrogen intermediates (Fig. 3A and figs. S27 and S28). The temperature of the reduction peak shifted toward lower temperatures with an increase in intimacy from M1 to M4. The peak position of the nanoscale-integrated 30FeV-3V/Al catalysts shifted to 475°C, which is near the position of the intrinsic reduction peak of 3V/Al and lower than that of $FeVO_4$. In situ XRD with temperatures ramping from 50° to 600°C under H_2/argon indicated that pure $FeVO_4$ NPs showed a phase transformation from triclinic phase $FeVO_4$ to spinel $Fe_{1.5}V_{1.5}O_4$ (PDF#04-005-6267) near 600°C. By contrast, the intensified hydrogen spillover facilitated the phase change of a $FeVO_4$ NP initially at 450°C when it was in proximity with 3V/Al (Fig. 3, B and C).

Analysis of the thermogravimetry relaxation curves showed a higher oxygen release rate for M4 than for $FeVO_4$ (fig. S29). Thus, the reactive hydrogen species from the dissociation of H_2 on VO_x sites may substantially expedite the reduction of adjacent $FeVO_4$ by hydrogen spillover. The combustion of hydrogen species on $FeVO_4$ can be considered as a "pump" for transferring hydrogen species. In the hydrogen-spillover effect identified here, hydrogen atoms were generated at the oxide site, in contrast to the conventional scenario, in which hydrogen atoms are primarily produced on the surface of metal NPs (*46*–*50*). Isothermal MS, in which propane was fed to 30FeV-3V/Al tandem catalysts at 550°C, showed the evolution of propylene and H_2 signals (fig. S30) along with an initial rise and downward depression in the H_2 signal and the appearance of H_2O signal. Thus, hydrogen was consumed through the hydrogen spillover–assisted SHC.

When cofeeding H_2 with C_3H_8 and C_3H_6, for pure $FeVO_4$, the formation rates of CO_x that resulted from the overoxidation of C_3H_8 and C_3H_6 were substantially reduced from 0.55 to 0.30 mmol g_{cat}^{-1} min^{-1} at 2 min on stream (fig. S31). Density functional theory (DFT) calculations evidenced the lower adsorption energy of H* compared with that of C_3H_7 and C_3H_6 on $FeVO_4$, and the increase in the H* coverage led to the decrease in the adsorption strength of C_3H_7 and C_3H_6 (fig. S32). Such results underpin a mechanism in which propane dissociated on the dispersed VO_x sites to produce hydrogen intermediates, which then spilled and preferentially oxidized over the adjacent $FeVO_4$ sites. This consequently suppressed the overoxidation of C_3H_6 (or other C_3H_8-derived species), with a distance-dependent hydrogen-spillover effect (see detailed discussion in fig. S32).

This hypothesis of hydrogen spillover–mediated coupling of PDH with CL-SHC was further supported by in situ diffuse reflectance infrared Fourier transform (DRIFT) experiments. By following the transient switches between the argon and C_3H_8 (10 vol % in argon) at 550°C, three distinct surface hydroxyl groups bands appeared at ~3750, 3690, and 3650 cm^{-1}, which we assigned to the Al_V–OH, Al_{IV}–OH, and V–OH (*23*, *51*), respectively (Fig. 3, D and E, and fig. S33). This finding was substantiated by hydrogen-deuterium (H-D) exchange experiments

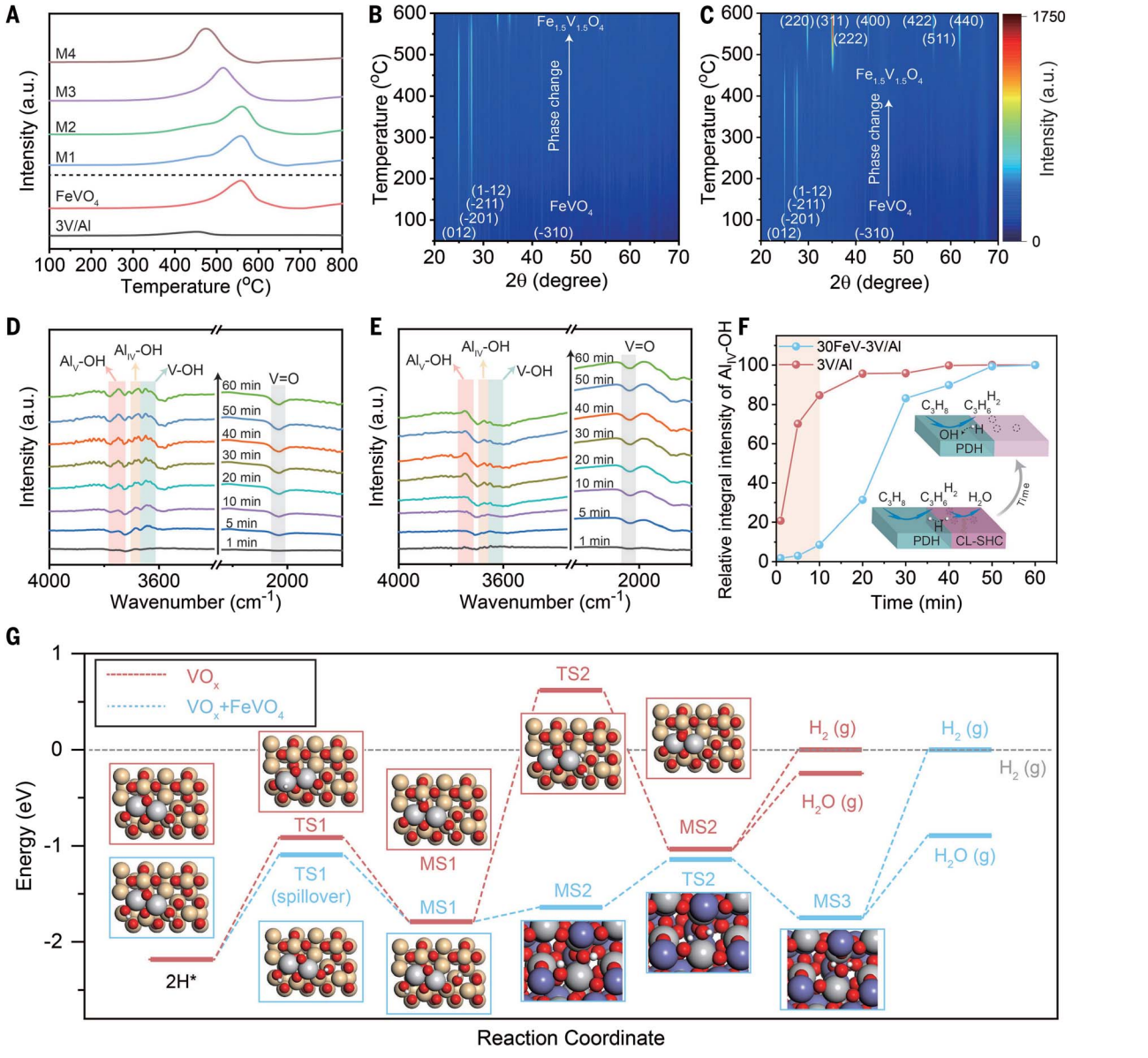


Fig. 3. Hydrogen spillover–mediated coupling of PDH and CL-SHC. (**A**) H_2-TPR of $FeVO_4$ and 3V/Al catalysts with different tandem modes. (**B** and **C**) In situ XRD patterns of $FeVO_4$ (B) and 30FeV-3V/Al (C) under 10 vol % H_2 ramping from 50° to 600°C. Numbers in parentheses indicate the representative crystalline planes. (**D** to **F**) In situ DRIFTs of 3V/Al (D) and 30FeV-3V/Al (E) and the relative intensity of Al_{IV}–OH (F) under 10 vol % C_3H_8 at a pressure of 1 torr at 550°C. The inset schematic in (F) depicts the evolution processes of Al_{IV}–OH. (**G**) Calculated potential energy diagrams of H* on VO_x/Al_2O_3 and VO_x/Al_2O_3+$FeVO_4$ with the corresponding structures along the reaction paths. Color scheme is as follows: O, red; Al, taupe; V, gray; and Fe, blue. TS indicates the transition state, and MS indicates the ending state, either H_2O(g) or H_2(g).

(fig. S34). The emergence of V–OH features indicated that C–H bond activation occurred on the VO_x site, accompanied by the reduction of the V=O bond. The characteristic peaks of Al_V–OH and Al_{IV}–OH were indicative of the cleavage of the V–O–Al bonds and the H species transferring to adjacent Al–O sites from the VO_x sites, respectively. Figure 3F shows the variations in integral intensities of the Al_{IV}–OH infrared signal under the C_3H_8 atmosphere with reaction time. On the 3V/Al catalysts, Al_{IV}–OH species experienced a sharp increase within a 10-min period, which accounts for almost 85% of the integral intensity of the infrared signal during the steady state. By contrast, Al_{IV}–OH species on the 30FeV-3V/Al catalysts were less dominant within the 10-min period and accounted for only 8% of integral intensity of the steady state.

We inferred that the adjacent $FeVO_4$ with nanoscale proximity triggered the migration of atomic hydrogen species from VO_x toward $FeVO_4$ to combust the hydrogen species. With the sustained consumption of lattice oxygen of $FeVO_4$, the driving force for hydrogen combustion becomes weak, and the hydrogen species diffuse to adjacent Al–O sites. This process was evidenced by the notable increase in Al_{IV}–OH species after the 10-min period. Similar results were observed on the 30FeV-3V/Al, 3V/Al, and $FeVO_4$ catalysts when exposed to the transient switches between the argon and H_2 (10 vol % in argon) conditions (fig. S35).

Theoretical calculations

DFT calculations offered a more comprehensive understanding of the adsorption and spillover characteristics of hydrogen on the surface of each catalyst. Hydrogen adsorption energies

(HAE) were first calculated on $FeVO_4$ and dimer VO_x on γ-Al_2O_3 (Fig. 3G and figs. S36 and S37). The process of hydrogen adsorption on $FeVO_4$ was exothermic, with an adsorption energy value (−1.49 eV) that was lower than those with V_2O_5/Al_2O_3 (−1.15 eV), V_2O_4/Al_2O_3 (−0.89 eV), and V_2O_3/Al_2O_3 (−0.86 eV). These findings suggest that the hydrogen spillover toward $FeVO_4$ becomes thermodynamically favorable in comparison with that toward VO_x/Al_2O_3. For dimer V_2O_4/Al_2O_3, the formation of H_2* (MS2 in red in Fig. 3G) had a high activation barrier. Therefore, the H* spillover (ΔE = 1.09 eV) took precedence over the formation of H_2O and H_2 (ΔE = 1.27 eV), illustrating that the occurrence of hydrogen spillover from the VO_x site to $FeVO_4$ is feasible. The reaction energy profile of the H intermediates toward H_2O (ΔE = 0.50 eV) indicated the favorable formation of H_2O(g) rather than H_2(g) on the $FeVO_4$ surface, denoting the selective H combustion on VO_x+$FeVO_4$ tandem catalysts (Fig. 3G).

Related tandem catalysts

To assess the generality and prowess of hydrogen spillover–assisted coupling of PDH and CL-SHC, we investigated the catalytic behaviors from two perspectives. We either altered the PDH catalysts by replacing 3V/Al with 3Zn/Al, 3Ga/Al, or 3Cr/Al or replaced $FeVO_4$ as the oxygen carrier with V_2O_5 in different tandem modes, as mentioned in the previous section of tandem catalyst systems (fig. S38). As expected, we observed strong correlations between the intimacy and selectivity of propylene (figs. S39 to S44), following "the closer, the better" criteria. Generally, the related tandem catalysts with nanoscale proximity exhibit better propylene yield than their corresponding parent systems. These results support that this hydrogen spillover–mediated coupling mechanism occurs in a tandem fashion with nanoscale proximity.

Conclusions

We demonstrated a hydrogen spillover–mediated coupling of PDH and CL-SHC at the nanoscale by using multifunctional tandem redox catalysts. Different tandem modes with site proximities ranging from the millimeter-scale to the nanoscale were investigated, validating the selectivity of propylene production in an intimacy-dependent manner. Experiments and theoretical analysis demonstrated that hydrogen intermediates formed at the VO_x sites migrated to adjacent $FeVO_4$ oxygen carriers for combustion, refining the tandem process of PDH with CL-SHC. Given that many industrial-relevant chemical conversion processes are endothermic and equilibrium-limited, we anticipate that this concept of hydrogen spillover–assisted CL-SHC and the coupling mechanism of tandem catalysis identified here will open new avenues to guide intensification of the chemical engineering process.

REFERENCES AND NOTES

1. J. J. H. B. Sattler, J. Ruiz-Martinez, E. Santillan-Jimenez, B. M. Weckhuysen, *Chem. Rev.* **114**, 10613–10653 (2014).
2. A. H. Motagamwala, R. Almallahi, J. Wortman, V. O. Igenegbai, S. Linic, *Science* **373**, 217–222 (2021).
3. D. Zhao *et al*., *Nature* **599**, 234–238 (2021).
4. R. Ryoo *et al*., *Nature* **585**, 221–224 (2020).
5. F. Cavani, N. Ballarini, A. Cericola, *Catal. Today* **127**, 113–131 (2007).
6. J. T. Grant *et al*., *Science* **354**, 1570–1573 (2016).
7. S. Vajda *et al*., *Nat. Mater.* **8**, 213–216 (2009).
8. H. Zhou *et al*., *Science* **372**, 76–80 (2021).
9. V. S. Bobrov, N. G. Digurov, V. V. Skudin, *J. Membr. Sci.* **253**, 233–242 (2005).
10. A. Schneemann *et al*., *Chem. Rev.* **118**, 10775–10839 (2018).
11. T. Waku, S. Y. Yu, E. Iglesia, *Ind. Eng. Chem. Res.* **42**, 3680–3689 (2003).
12. H. Yan *et al*., *Science* **371**, 1257–1260 (2021).
13. L. Zeng, Z. Cheng, J. A. Fan, L. S. Fan, J. L. Gong, *Nat. Rev. Chem.* **2**, 349–364 (2018).
14. R. W. Breault, Eds., *Handbook of Chemical Looping Technology* (Wiley, 2018).
15. X. Zhu, Q. Imtiaz, F. Donat, C. R. Muller, F. X. Li, *Energy Environ. Sci.* **13**, 772–804 (2020).
16. R. K. Grasselli, D. L. Stern, J. G. Tsikoyiannis, *Appl. Catal. A* **189**, 9–14 (1999).
17. L. Låte, W. Thelin, E. A. Blekkan, *Appl. Catal. A* **262**, 63–68 (2004).
18. J. G. Tsikoyiannis, D. L. Stern, R. K. Grasselli, *J. Catal.* **184**, 77–86 (1999).
19. S. Kim, L. Annamalai, R. F. Lobo, *Chem. Eng. J.* **455**, 140919 (2023).
20. R. B. Dudek, F. Li, *Fuel Process. Technol.* **218**, 106827 (2021).
21. G. Rothenberg, E. A. de Graaf, A. Bliek, *Angew. Chem. Int. Ed.* **42**, 3366–3368 (2003).
22. K. Routray, W. Zhou, C. J. Kiely, I. E. Wachs, *ACS Catal.* **1**, 54–66 (2011).
23. Z.-J. Zhao *et al*., *Angew. Chem. Int. Ed.* **57**, 6791–6795 (2018).
24. M. Ek, Q. M. Ramasse, L. Arnarson, P. Georg Moses, S. Helveg, *Nat. Commun.* **8**, 305 (2017).
25. G. Mitran *et al*., *Catal. Today* **306**, 260–267 (2018).
26. S. K. Matam *et al*., *Catalysts* **10**, 1370 (2020).
27. X. Zhang, Y. Yue, Z. Gao, *Catal. Lett.* **83**, 19–25 (2002).
28. J.-L. Sánchez-García *et al*., *Top. Catal.* **65**, 1408–1418 (2022).
29. Y. Lei *et al*., *Chem. Mater.* **24**, 3525–3533 (2012).
30. M. M. Hossain, *Ind. Eng. Chem. Res.* **56**, 4309–4318 (2017).
31. S. Chen *et al*., *J. Am. Chem. Soc.* **141**, 18653–18657 (2019).
32. S. Crapanzano, I. V. Babich, L. Lefferts, *Appl. Catal. A* **391**, 70–77 (2011).
33. S. Chen *et al*., *Angew. Chem. Int. Ed.* **59**, 22072–22079 (2020).
34. S. Rostom, H. I. de Lasa, *Ind. Eng. Chem. Res.* **56**, 13109–13124 (2017).
35. R. K. Grasselli, D. L. Stern, J. G. Tsikoyiannis, *Appl. Catal. A* **189**, 1–8 (1999).
36. M. P. Lobera, C. Tellez, J. Herguido, M. Menendez, *Ind. Eng. Chem. Res.* **47**, 9314–9320 (2008).
37. S. Kaneko, T. Arakawa, M. Ohshima, H. Kurokawa, H. Miura, *Appl. Catal. A* **356**, 80–87 (2009).
38. G. Sun *et al*., *Nat. Commun.* **9**, 4454 (2018).
39. W. Cai *et al*., *Sci. Adv.* **4**, eaar5418 (2018).
40. F. Xing, J. Ma, K. I. Shimizu, S. Furukawa, *Nat. Commun.* **13**, 5065 (2022).
41. C. P. Robert, G. Casella, *Monte Carlo Statistical Methods* (Springer, 1999).
42. A. M. Johansen, *International Encyclopedia of Education* (Elsevier, ed. 3, 2010), pp. 296–303.
43. P. B. Weisz, E. W. Swegler, *Science* **126**, 31–32 (1957).
44. H. Pines, W. O. Haag, *J. Am. Chem. Soc.* **82**, 2471–2483 (1960).
45. X. Wang, I. E. Wachs, *Catal. Today* **96**, 211–222 (2004).
46. W. Karim *et al*., *Nature* **541**, 68–71 (2017).
47. R. Prins, *Chem. Rev.* **112**, 2714–2738 (2012).
48. L. Jiang *et al*., *Nat. Nanotechnol.* **15**, 848–853 (2020).
49. M. S. Frei *et al*., *Nat. Commun.* **10**, 3377 (2019).
50. M. Xiong, Z. Gao, Y. Qin, *ACS Catal.* **11**, 3159–3172 (2021).
51. G. Busca, *Catal. Today* **226**, 2–13 (2014).

ACKNOWLEDGMENTS

We thank D. Fu for fruitful discussions. We also thank two anonymous reviewers for constructive comments. W.W. and R.L. thank the National University of Singapore. **Funding:** This work was supported by the National Key R&D Program of China (2021YFA1501302), the National Science Foundation of China (nos. 22121004, 22122808, and U1862207), the Haihe Laboratory of Sustainable Chemical Transformations, the Program of Introducing Talents of Discipline to Universities (BP0618007), and the XPLORER PRIZE. **Author contributions:** J.G. initiated and supervised the research. W.W., J.S., and H.S. performed the catalyst preparation. W.W., S.C., and J.S. performed the catalytic reaction test. W.W., S.C., G.S., and X.W. performed the characterization. R.L., Z.-J.Z., and J.G. contributed to DFT calculations. W.W., S.C., C.P., Z.-J.Z., and J.G. designed the study, analyzed the data, and wrote the paper. All authors contributed to the discussion of the research and approved the final version of the manuscript. **Competing interests:** J.G., W.W., S.C., C.P., and Z.-J.Z. are inventors listed on Chinese patents (202310248826.9 and 202310248877.1, filing date 15 March 2023, pending) filed by Tianjin University, which cover the PDH+CL-SHC process and tandem catalysts reported in this paper. **Data and materials availability:** All data are available in the main text or the supplementary materials. **License information:** Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adi3416
Materials and Methods
Figs. S1 to S44
Tables S1 to S4
References (*52*–*63*)

Submitted 19 April 2023; accepted 13 July 2023
Published online 27 July 2023
10.1126/science.adi3416

PLANT SCIENCE

Functional diversification of a wild potato immune receptor at its center of origin

Yerisf C. Torres Ascurra[1], Lisha Zhang[2], AmirAli Toghani[3], Chenlei Hua[2], Nandeesh Jalahalli Rangegowda[1], Andres Posbeyikian[3], Hsuan Pai[3], Xiao Lin[1]†, Pieter J. Wolters[1], Doret Wouters[1], Reinhoud de Blok[1], Niels Steigenga[1]‡, Maxence J. M. Paillart[4], Richard G. F. Visser[1], Sophien Kamoun[3], Thorsten Nürnberger[2,5]*, Vivianne G. A. A. Vleeshouwers[1]*

Plant cell surface pattern recognition receptors (PRRs) and intracellular immune receptors cooperate to provide immunity to microbial infection. Both receptor families have coevolved at an accelerated rate, but the evolution and diversification of PRRs is poorly understood. We have isolated potato surface receptor Pep-13 receptor unit (PERU) that senses Pep-13, a conserved immunogenic peptide pattern from plant pathogenic *Phytophthora* species. PERU, a leucine-rich repeat receptor kinase, is a bona fide PRR that binds Pep-13 and enhances immunity to *Phytophthora infestans* infection. Diversification in ligand binding specificities of PERU can be traced to sympatric wild tuber-bearing *Solanum* populations in the Central Andes. Our study reveals the evolution of cell surface immune receptor alleles in wild potato populations that recognize ligand variants not recognized by others.

Plant cell surface pattern recognition receptors (PRRs) and intracellular immune receptors cooperate to provide robust resistance to microbial infection (*1*–*3*). The synergistic activation of plant immunity by spatially separated plant immune receptors suggests their coevolution. A strong correlation in the number of genes encoding surface and intracellular immune receptors observed across the plant lineage supports the concept of mutual potentiation of immune responses initiated in different plant cell compartments (*4*, *5*).

It is assumed that pathogen pressure in ecological niches drives plant immune receptor evolution (*5*, *6*), but evidence for diversification of plant PRR sequences and functions among natural plant populations is lacking. We hypothesized that a PRR might recognize Pep-13, a conserved microbial immunogenic 13-amino-acid fragment from a cell wall glycoprotein (GP42) with transglutaminase (TG) activity (*7*–*9*). TGs are produced by several plant-pathogenic oomycete *Phytophthora* species, including *P. infestans*, the causal agent of potato late blight disease and the Great Irish Famine (*10*–*12*). Pep-13 triggers a hypersensitive response and other immunity-associated responses in diverse plant species, including the solanaceous host plant, potato (*10*).

[1]Plant Breeding, Wageningen University and Research, 6708 PB Wageningen, Netherlands. [2]Department of Plant Biochemistry, Centre of Plant Molecular Biology (ZMBP), University of Tübingen, 72076 Tübingen, Germany. [3]The Sainsbury Laboratory, University of East Anglia, Norwich, UK. [4]Wageningen Food & Biobased Research, Wageningen University and Research, 6708 WG Wageningen, Netherlands. [5]Department of Biochemistry, University of Johannesburg, Johannesburg 2006, South Africa.
*Corresponding authors. Email: vivianne.vleeshouwers@wur.nl (V.G.A.A.V.); nuernberger@uni-tuebingen.de (T.N.)
†Present address: State Key Laboratory of Plant Genomics, Institute of Microbiology, Chinese Academy of Sciences, Beijing 100101, China. ‡Present address: Department of Agronomy, Purdue University, West Lafayette, IN 47906, USA.

Potato *PERU* senses oomycete-derived pattern Pep-13

We screened a collection of wild *Solanum* species and cultivated potato genotypes for cell death induction when infiltrated with Pep-13 or its structural derivative, Pep-25 (*7*, *13*). To identify the Pep-13-receptor by a map-based cloning approach, we crossed genotype *Solanum tuberosum* Group Phureja DM 1-3 516 R44 (DM) and genotype *S. tuberosum* RH89-039-16 (RH) (Fig. 1A). DM is a Pep-13/Pep-25–sensitive genotype, which was previously used to establish the potato reference genome (*14*). RH is a Pep-13/Pep-25–insensitive genotype. We back-crossed the F1 generation 3240-4 to the RH parent and used the resulting F2 population (3648) for genetic mapping (*15*). Pattern sensitivity segregated in a 1:1 ratio, suggesting that a single, dominant gene encodes the corresponding receptor (fig. S1A). Pep-13/Pep-25 sensitivity was previously mapped to the top of chromosome 3 (*13*), and subsequent marker-assisted fine-mapping yielded a 55.2-kb fragment containing 7 open reading frames, three of which encode leucine-rich repeat receptor kinases (LRR-RKs a-c) (fig. S1B). LRR-RKs consist of an extracellular LRR domain, a transmembrane-spanning domain, and an intracellular serine/threonine protein kinase domain, which is absent in LRR receptor proteins (LRR-RPs). LRR ectodomain-containing receptors are the predominant type of plant PRRs known to date and have evolved to recognize primarily proteinaceous microbial patterns or phytocytokines (*16*–*18*).

To determine which LRR-RK candidate gene sequence confers Pep-13 sensitivity, we performed transient expression assays in the solanaceous model plant, *Nicotiana benthamiana*. *Agrobacterium* infection–mediated expression of LRR-RK b—but not LRR-RK a or c-enco cDNA sequences—resulted in plant cell d after treatment with either Pep-13 or GP42 (Fig. 1B). We thus designated LRR-RK b Pep-13 receptor unit (PERU). PERU is a canonical plant LRR-RK that hosts an ectodomain composed of 27 LRRs linked by a transmembrane domain to an intracellular serine-threonine protein kinase domain (fig. S2). Stable expression of *PERU* cDNA in Pep-13–insensitive potato cultivar Atlantic resulted in Pep-13–inducible cell death, production of reactive oxygen species (ROS) and accumulation of the plant stress hormone, ethylene (Fig. 1, C to E, and fig. S3). These responses were not observed in wild-type (WT) Atlantic or in control lines transformed with empty vector only. Inactivation of the *PERU* locus in Pep-13–sensitive DM by CRISPR-Cas9 mutagenesis provided direct proof for a causal role of PERU in Pep-13 pattern recognition. To abolish *PERU* gene expression, genotype DM was stably transformed with CRISPR-Cas9 and 4 sgRNA, CRISPR lines were genotyped, and deletion and frameshift mutations were found, resulting in loss of Pep-13–induced cell death, ROS burst, and ethylene production (Fig. 1, F to H). In sum, these results document a role for potato PERU in Pep-13 pattern recognition.

A PERU-SERK3 complex mediates Pep-13/Pep-25–induced defenses and plant cell death

LRR-type PRRs recognize their cognate ligands by binding to their LRR ectodomains (*19*). We investigated ligand-receptor binding in vitro and in planta. To test for physical interaction of PERU and Pep-25 in vitro, we incubated recombinant hexa-histidine (His_6)–tagged PERU LRR ectodomain protein ($PERU^{LRR}$-His_6) with biotinylated Pep-25 (Pep-25-bio) before treatment with the homo-bifunctional cross-linker ethylene glycol bis (succinimidyl succinate) (EGS) to stabilize the ligand-receptor complex (*8*). Pep-25-bio is as active as Pep-25 (fig. S4). Following $PERU^{LRR}$-His_6 immunoprecipitation, bound Pep-25-bio was visualized by streptavidine/anti-streptavidine antisera (fig. S5). A large molar excess of free Pep-13 competitively abolished ligand-receptor complex formation, which suggests direct and specific ligand binding by $PERU^{LRR}$. The affinity constant of the ligand/receptor interaction (K_D = 88.9 nM) is close to ligand concentrations required for immune activation in *p35S::PERU-expressing* Pep-13–insensitive *Solanum hjertingii* or in Pep-13–sensitive potato DM (EC_{50} = 9.8 nM or 44 nM respectively), indicating that the PERU ectodomain is sufficient for ligand binding (fig. S5). To analyze the Pep-25-bio/PERU interaction in planta, we treated leaves of *N. benthamiana* plants expressing green fluorescent protein (GFP)–tagged PERU (*p35S::PERU-GFP)* with ligand prior to EGS treatment. Precipitation of PERU-GFP protein and subsequent detection of bound ligand corroborated

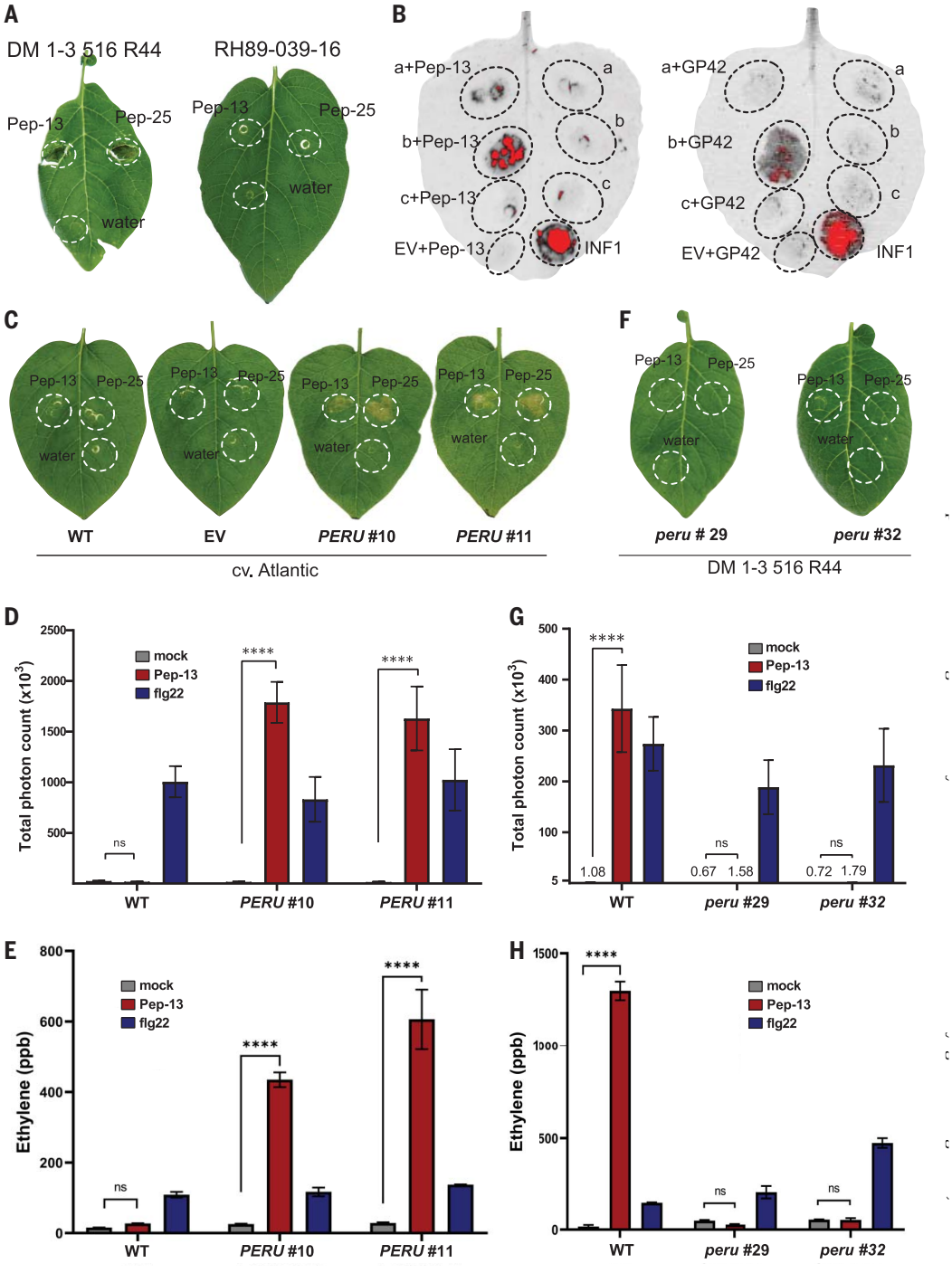


Fig. 1. PERU confers response to Pep-13. (**A**) The genotype DM 1-3 516 R44 (DM) shows cell death response upon infiltration of Pep-13 and Pep-25 whereas the genotype RH89-039-16 (RH) does not. (**B**) Representative *N. benthamiana* leaves co-agroinfiltrated with the candidate genes "a," "b," or "c," and Pep-13 or the full-length glycoprotein GP42, show that candidate "b" confers cell death to Pep-13 (left leaf) and GP42 (right leaf), whereas candidates "a" and "c" do not. Cell death was visualized by a red light imaging system at 3 days post infiltration (*47*). (**C**) Potato cultivar Atlantic (WT) is insensitive to Pep-13/25, transgenic Atlantic expressing *PERU* (*PERU* #10 and *PERU* #11) show cell death after Pep-13/25 infiltration, cv. Atlantic transformed with empty vector (EV) is included as negative control. (**D**) Total ROS production, and (**E**) Ethylene accumulation after treatment with 1 μM Pep-13, flg22 or water as control in potato cv. Atlantic WT, EV, *PERU* #10 and #11. (**F**) CRISPR lines *peru* #29 and #32 are insensitive to Pep-13/25. (**G**) Total ROS production, and (**H**) Ethylene accumulation after treatment with 1 μM Pep-13, flg22, or water as control, in DM (WT) and lines *peru* #29 and *peru* #32. Data were analyzed using one-way analysis of variance (ANOVA) with Tukey's test, (****P-value ≤ 0.0001). All experiments were performed three times with similar results and representative experiments are shown.

ligand-receptor binding observed in vitro (Fig. 2A). Again, excess of Pep-13 abolished ligand binding. We did not observe an inhibitory effect when a biologically inactive Pep-13 mutant peptide, Pep-13W231A (tryptophan residue 231 mutated to alanine, amino acid numbering corresponds to full-length GP42 sequence) (*10*), was used as a competitor (Fig. 2A). Notably, a W231A mutant of GP42 not only abolished its plant defense-eliciting activity but also reduced its TG activity by 98.5% (*10*). Altogether, these data demonstrate specific binding of Pep-13 to its high-affinity binding site, PERU.

The LRR-RK BAK1/SERK3 (BRASSINOSTEROID INSENSITIVE 1-ASSOCIATED KINASE1/SOMATIC EMBRYOGENESIS RECEPTOR KINASE 3) forms ligand-induced receptor/co-receptor complexes with plant LRR-RK-type PRRs (*17*, *18*, *20*). We found Pep-13 pattern-induced complex formation of PERU and SERK3A after transient coexpression of *p35S::PERU^{DM}-GFP* and *p35S::SERK3A^{DM}-Myc* in *N. benthamiana* plants (Fig. 2B). Virus-induced gene silencing of *SERK3* (*TRV::NbSERK3*) in *N. benthamiana* resulted in a massive reduction in Pep-13–induced hypersensitive cell death and ROS production in *TRV::NbSERK3* plants (Fig. 2C and fig. S6), suggesting that PERU recruits SERK3 in a pattern-dependent manner. Silencing of SOBIR1 (SUPPRESSOR OF BAK1-INTERACTING KINASE1) (*TRV::NbSOBIR1/-like*), which is exclusively required for the function of LRR-RP-type PRRs, did not

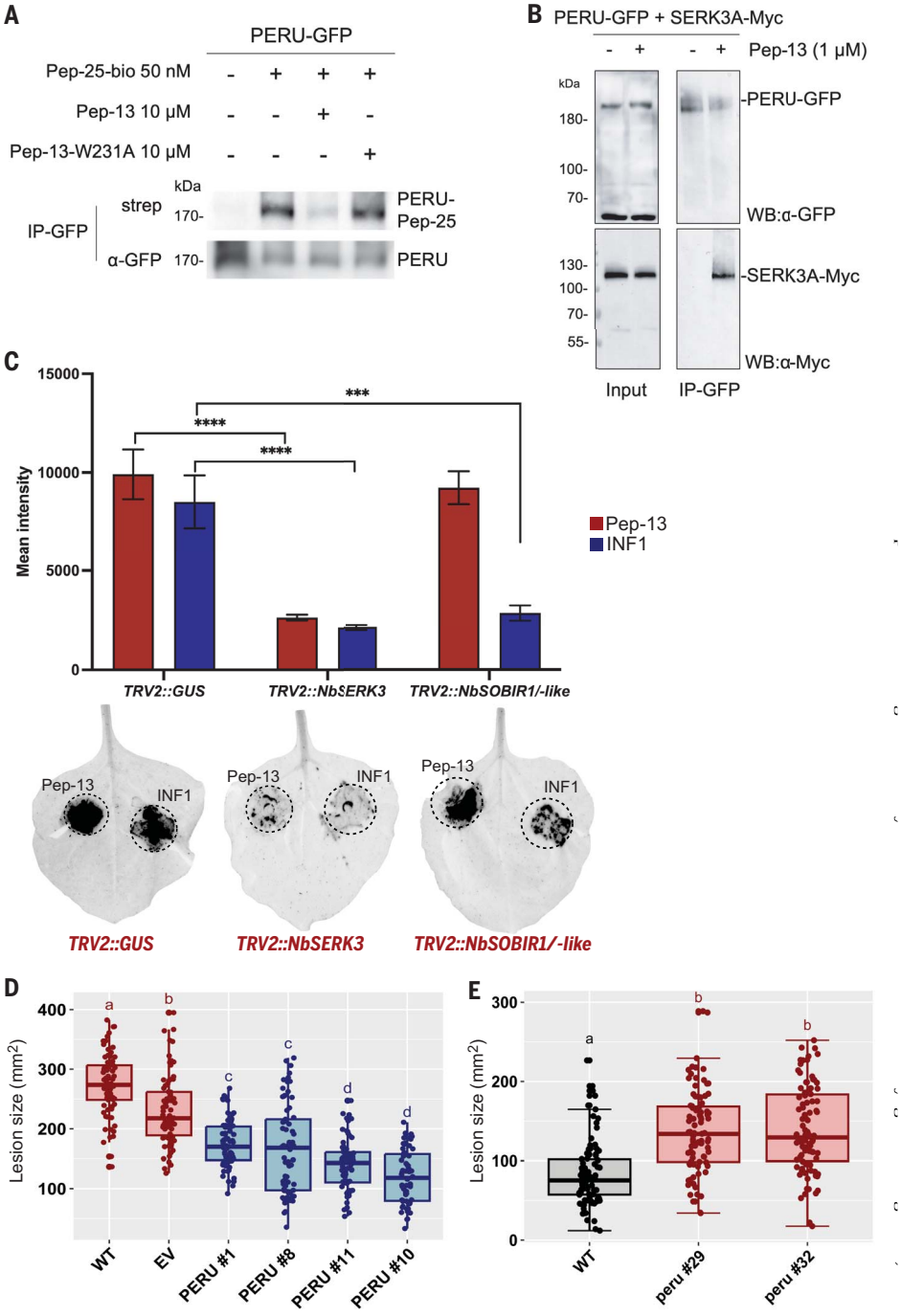


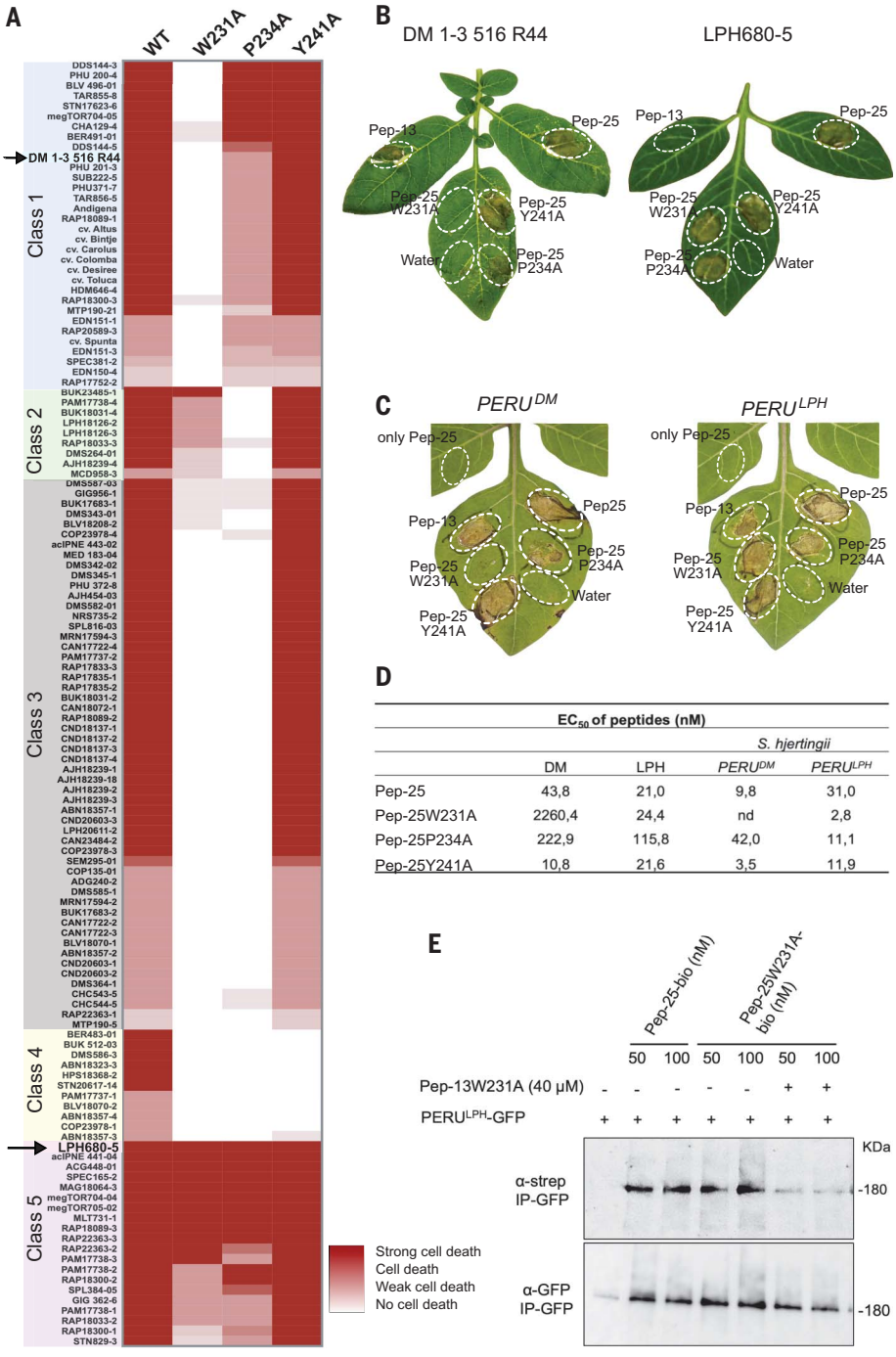
Fig. 2. PERU recognizes Pep-13, is dependent on SERK3, and confers enhanced resistance to *Phytophthora infestans*. (**A**) Binding of biotinylated Pep-25 to PERU. PERU-GFP transiently expressed in *N. benthamiana* served as receptor, Pep-25-bio peptide as ligand, and unlabeled Pep-13 and Pep-13W231A peptides as competitors. Streptavidin-AP (strep) was used to detect ligand binding to receptor. (**B**) PERU-GFP and SERK3A^{DM}-Myc were transiently expressed in *N. benthamiana* and treated with Pep-13 or water as control. Proteins were subjected to coimmunoprecipitation with GFP trap beads and immunoblotting with tag-specific antibodies. This experiment was performed in duplicate with similar results. (**C**) Cell death induced by Pep-13 peptide and INF1 protein (as control) in TRV2::*NbSERK3 N. benthamiana* leaves was significantly reduced compared with TRV2::*GUS*. Pep-13–induced cell death was not reduced in leaves treated with TRV2::*Nb*SOBIR1/-like. Representative leaves are shown at the bottom. Data were analyzed using one-way ANOVA with Tukey's test, (****P-value ≤ 0.0001, ***P-value < 0.001, **P-value < 0.01). (**D**) Leaves of cv. Atlantic WT, and transgenic EV (negative control), *PERU* #1, #8, #10, and #11 intact plants were spot-inoculated with *P. infestans* strain Dinteloord, and lesion sizes measured at 4 dpi. All transgenic lines expressing *PERU* showed smaller lesions than WT or EV. (**E**) DM (WT) and CRISPR lines *peru* #29 and #32 were spot-inoculated with *P. infestans* strain Dinteloord. Lesion sizes were measured at 4 dpi. Larger lesions were observed in the CRISPR lines *peru* #29 and #32. Data were analyzed using one-way ANOVA with Tukey's test (P-value ≤ 0.05), different letters indicate significant differences. All experiments were performed 3 times with similar results and representative results are shown.

affect Pep-13–induced cell death formation. By contrast, *Phytophthora infestans* elicitin INF1–induced cell death mediated by activation of LRR-RP-type ELR (ELICITIN RESPONSE) (*21*) is reduced in both TRV::Nb*SERK3* and TRV::Nb*SOBIR1/-like* plants (Fig. 2C).

In solanaceous *N. benthamiana*, activation of plant immunity and cell death by LRR-RP-type PRRs requires lipase-like ENHANCED DISEASE SUSCEPTIBILITY 1 (EDS1) (*22*) and helper NUCLEOTIDE-BINDING LRR (hNLR) REQUIRED FOR HYPERSENSITIVE RESPONSE-ASSOCIATED CELL DEATH 2, 3, and 4 (NRC2, NRC3, NRC4) (*23*, *24*). Because LRR-RK-type PRRs have not previously been implicated in activating plant cell death in any plant system, we tested whether these proteins are required for PERU signaling. *N. benthamiana* plants transiently expressing *p35S::PERU*^{DM} developed cell death symptoms upon infiltration of Pep-13 or GP42 (Fig. 1B) and produced ethylene upon treatment with Pep-25 (fig. S4). Transient expression of *p35S::PERU*^{DM}*-GFP* in *N. benthamiana* mutants lacking EDS1 and related PHYTOALEXIN-DEFICIENT 4 (PAD4) (*25*, *26*) or hNLRs NRC2/3/4 (*23*) had no reducing effect on elicitor–induced cell death and ethylene production (fig. S7). Altogether, we find substantial differences in the molecular mechanisms controlling plant immune responses upon activation of different classes of LRR-type PRRs in this plant.

PERU confers enhanced resistance to *P. infestans*

Potato varieties used in agricultural production are often susceptible to major plant pathogens,



D

EC_{50} of peptides (nM)				
			S. hjertingii	
	DM	LPH	$PERU^{DM}$	$PERU^{LPH}$
Pep-25	43,8	21,0	9,8	31,0
Pep-25W231A	2260,4	24,4	nd	2,8
Pep-25P234A	222,9	115,8	42,0	11,1
Pep-25Y241A	10,8	21,6	3,5	11,9

Fig. 3. Functional diversification of PERU in *Solanum* section *Petota*. (**A**) Heat map of cell death responses of *Solanum* accessions to Pep-25 (WT) and Pep-25W231A, Pep-25P234A, Pep-25Y241A mutants reveals five different classes (1 to 5) of recognition specificities. DM 1-3 516 R44 (class 1) and LPH680-5 (class 5) are marked with arrows. (**B**) Cell death induction by Pep-13, Pep-25, and mutants in DM and LPH plants. In DM, Pep-25W231A induces no cell death, and Pep-25P234A induces weaker cell death whereas in LPH, Pep-25W231A and Pep-25-P234A-induced cell death is similar to that caused by Pep-25. (**C**) $PERU^{DM}$ or $PERU^{LPH}$ were transiently expressed in Pep-13–insensitive *S. hjertingii* by agroinfiltration. Infiltration of Pep-13/25 and Pep-25 mutant peptides yielded the same response pattern as observed in DM and LPH plants. (**D**) EC_{50} values were determined by quantification of elicitor-induced production of ethylene in *DM* or LPH plants, and in *S. hjertingii* plants transiently transformed with either $PERU^{DM}$ or $PERU^{LPH}$; nd, not determined. (**E**) Receptor/ligand binding assays show that $PERU^{LPH}$ specifically bound both Pep-25-bio and Pep-25W231A-bio as ligands, and Pep-13W231A efficiently competes for ligand binding to $PERU^{LPH}$. All experiments were performed 3 times with similar results and representative experiments are shown.

including species of the genus *Phytophthora* (*27*). Genetic engineering provides one way of increasing plant resistance in crop plants. Ectopic expression of plant PRRs is known to confer novel microbial pattern recognition specificities and enhanced pathogen resistance to crop plants (*16*, *21*, *28*, *29*). For infection assays with the virulent *P. infestans* isolate, Dinteloord, we obtained four potato cultivar Atlantic lines stably expressing *p35S::PERU*, and two DM lines stably transformed with CRISPR-Cas9 and 4 sgRNA to disable *PERU* gene expression. By scoring disease lesions four days post infection we found that *PERU*-transgenic lines were significantly less damaged when compared with WT plants or to lines transformed with vector only (Fig. 2D). Likewise, CRISPR-Cas9–generated *peru* mutant lines were significantly more susceptible to infection when compared with WT DM (Fig. 2E). Hence, PERU confers quantitative resistance against the pervasive potato late blight pathogen.

Diversification of PERU ligand specificities in wild potato populations

Pep-13/25–induced plant defenses have been studied in parsley and potato cell suspensions, as well as in leaves of a cultivated potato clone, Désirée (*7*, *10*). Alanine scanning mutagenesis of Pep-13 sequences revealed that mutant Pep-13W231A abolished elicitor activity, mutation of proline 234 (Pep-13P234A) reduced it, and replacement of the remaining amino acid residues (including tyrosine 241, Pep-13Y241A) did not significantly affect activities of the mutant peptides (*7*, *8*, *10*). We found the same pattern of ligand responses in the Pep-25–sensitive genotype DM, which was used for PERU identification (Figs. 1A, 3A, and data S1). To determine the frequency of biologically active *PERU* alleles in *Solanum* sect. *Petota* (*30*), we analyzed 476 genotypes corresponding to 98 species (97 wild, and 1 cultivated potato species) for cell death triggered by Pep-25 and its described mutant variants (data S1). 350 (74%) of these genotypes did not develop cell death in response to Pep-25, indicating that most wild *Solanum* genotypes in this collection lack an active *PERU* allele (data S1). Pep-25 and its mutants were tested for cell death induction on all 126 Pep-25–sensitive genotypes (Fig. 3A). Overall, we observed at least five different recognition specificities, including substantial qualitative and quantitative variations in their abilities to respond to Pep-25 and its mutants. Wild potato genotypes grouped in class 1 include DM and exhibit the same ligand response patterns as described for cultivated potato cultivar Désirée previously (*7*, *8*, *10*) (Fig. 3A). These accounted for 25% of all Pep-25–sensitive genotypes (Fig. 3A). Other genotypes showed responsiveness to Pep-13W231A but failed to respond to Pep-13P234A (class 2), others failed to respond to both Pep-13P234A and Pep-13W241A (class 3),

and some genotypes failed to respond to all mutant peptides of Pep13 (class 4). Some *Solanum* genotypes, such as *Solanum leptophyes* (LPH) 680-5, exhibit sensitivities to all Pep-25 variants tested and were categorized as class 5 genotypes (Fig. 3A). Altogether, our findings reveal that wild *Solanum* species bear diverse *PERU* alleles that differ from $PERU^{DM}$ and thus encode PRRs with distinct ligand specificities.

We assessed plant defense and cell death–inducing activities of Pep-25 WT and mutant peptides in potato genotypes expressing $PERU^{DM}$ (class 1 genotype) or $PERU^{LPH}$ (class 5 genotype) (Fig. 3B) in greater detail. We found that $PERU^{LPH}$-expressing plants recognized all Pep-25 variants, whereas $PERU^{DM}$ did not mount cell death in response to Pep-25W231A. To corroborate these findings and to rule out that potato

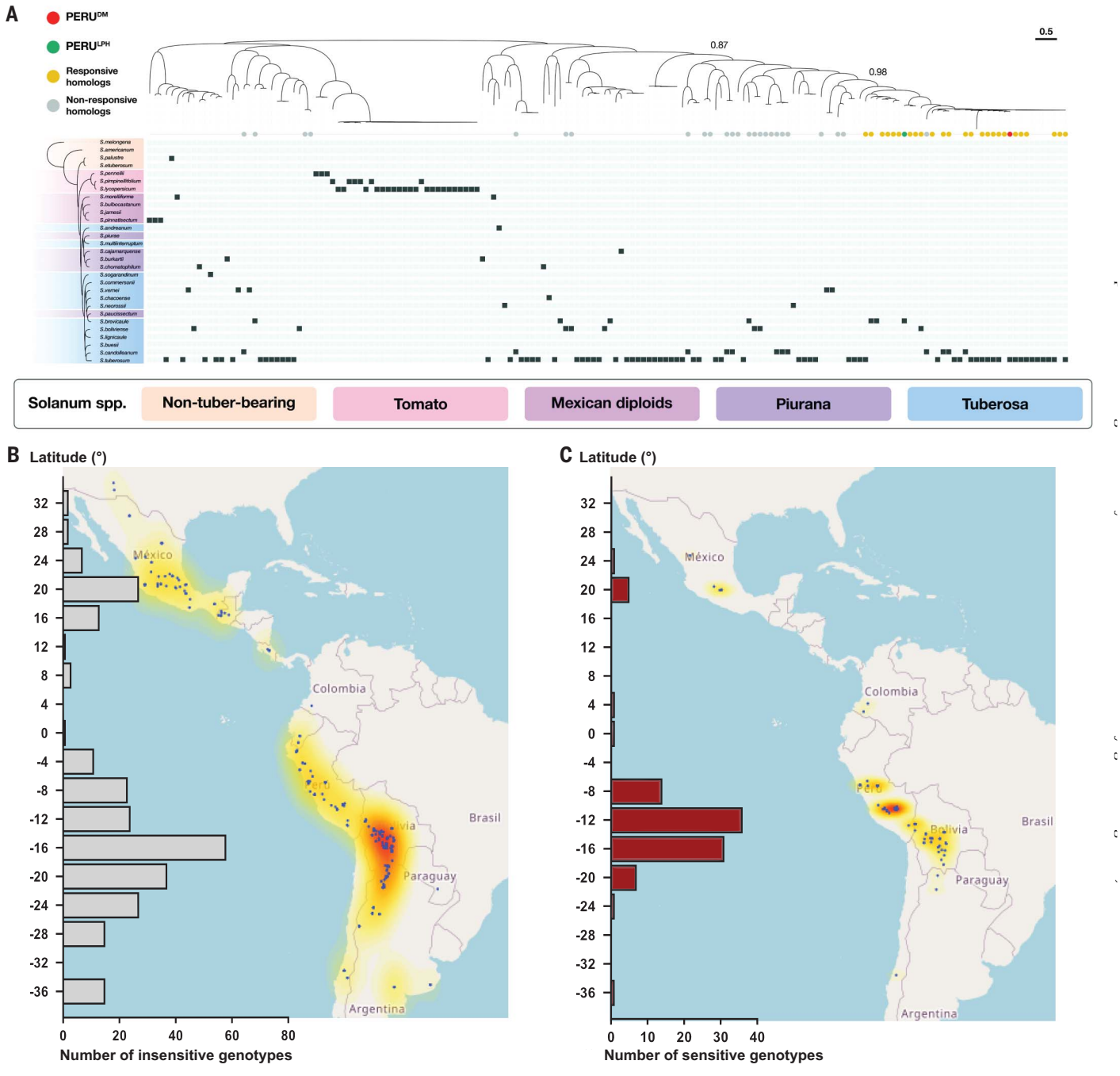


Fig. 4. Phylogenetics and geographic distribution of PERU. (**A**) Heatmap representation of presence/absence of PERU homologs across *Solanum* species. The $PERU^{DM}$, $PERU^{LPH}$, and the Pep-13 responsive homologs grouped together in a single clade whereas the nonresponsive homologs were distributed throughout the phylogenetic tree. Bootstrap values are shown for the clade containing all tested homologs (0.87) and the clade containing the responsive $PERU^{DM}$ homologs (0.98). The *Solanum* phylogeny was adapted from (*48*). (**B**) Kernel density distribution map of 266 genotypes insensitive to Pep-25 that are distributed from the Southern USA to Northern Chile; and (**C**) Kernel density distribution map of 98 sensitive genotypes that cluster in Peru and Bolivia. Red shades indicate high density, yellow shades indicate lower density, the blue dots represent individual geo-coordinates of accessions, and bar plots represent the number of genotypes along different latitudes of the continent. Available geographic coordinates of 364 genotypes (data S1) were used to elaborate the maps.

genotype–specific properties account for altered PERU ligand binding specificities, we transiently expressed $PERU^{DM}$ or $PERU^{LPH}$–encoding sequences in Pep-13–insensitive *S. hjertingii*. Again, infiltration of Pep-25 WT and mutant peptides yielded the same response pattern as observed previously (Fig. 3B), with all Pep-25 mutants inducing cell death in $PERU^{LPH}$ plants only (Fig. 3C). Thus, differences in ligand specificities of PERU proteins from $PERU^{DM}$ or $PERU^{LPH}$-expressing plants are features of the receptor proteins themselves rather than of co-receptors or other auxiliary factors.

We further analyzed biological activities of Pep-25 and its mutants by quantifying elicitor-induced production of the stress hormone ethylene in $PERU^{DM}$ or $PERU^{LPH}$ potato plants and in *S. hjertingii* plants transiently transformed with either *PERU* allele (Fig. 3, D and E). Determination of elicitor concentrations required to induce half-maximal ethylene production (EC_{50}) corroborated qualitative data from cell death assays (Fig. 3, B and C). Pep-25W231A proved as active as Pep-25 only when tested on $PERU^{LPH}$-expressing plants. We found substantially reduced or no activity of this peptide in plants expressing $PERU^{DM}$ (Fig. 3, D and E). In agreement with that, $PERU^{LPH}$ plants bound both Pep-25-bio and Pep-25W231A-bio in receptor-ligand binding assays (Fig. 3E). Likewise, Pep-13W231A efficiently blocked ligand binding to $PERU^{LPH}$ (Fig. 3E), but not to $PERU^{DM}$ (Fig. 2A). Altogether, our data obtained from ligand binding assays and from plant defense activation studies confirm that $PERU^{DM}$ and $PERU^{LPH}$ encode related LRR-RK immune receptors that have diversified in ligand specificities. Our findings further suggest that functional diversification has occurred within this immune receptor family during evolution, resulting in PERU alleles that recognize Pep-13 variants not recognized by others.

Evolutionary history of PERU

To obtain information about the origin of *PERU* alleles, we studied their geographic distribution and genetic variation (Fig. 4). We developed a computational pipeline to extract PERU sequences from 6,630,292 predicted proteins from 124 Solanaceae genome assemblies and for comparison extracted sequences of FLS2 (FLAGELLIN-SENSING 2), a conserved LRR-RK that detects bacterial flagellin (fig. S8 and data S2) (*3*). Both PERU and FLS2 clustered in well-supported clades within the LRR-RK subgroup XII indicating a monophyletic origin within the Solanaceae (fig. S9) (*5*). All plant LRR-RK-type PRRs currently known fall into this clade, including *Arabidopsis* FLS2 and EF-TU RECEPTOR (EFR), or rice Xa21 (*16*, *31*). This pipeline yielded 114 PERU clade sequences from 17 species and 180 FLS2 clade sequences from 33 species with the PERU clade sequences exhibiting markedly more diversity than the FLS2 clade (fig. S9 and data S3, S4, and S5). The same primers employed to isolate *PERU* (DM) facilitated the isolation of 26 responsive homologs and 25 nonresponsive homologs (data S6). The amplified and genome-extracted sequence datasets were then combined for phylogenomic analyses, which revealed that the PERU sequences encoding Pep-13–responsive PERU alleles fall into one clade whereas the nonresponsive homologs are scattered throughout the tree (Fig. 4A and fig. S10, A and B). The $PERU^{LPH}$ sequences are embedded within a tighter $PERU^{DM}$ clade indicating that evolution of a new ligand specificity and, hence, functional diversification has occurred within the PERU receptor family of *Solanum* (Fig. 4A and fig. S10A). Our phylogenomics analyses of *PERU* sequences further suggest that potato $PERU^{DM}$ and the PEP-13 receptor in parsley are distinct proteins, although they share similar ligand specificities (Fig. 3) (*7*, *10*).

Metazoan and plant immune receptors have been targeted by positive, diversifying selection, which accelerates the divergence between homologous proteins (*32*, *33*). To identify amino acids under diversifying selection in the proteins encoded by *PERU* alleles, we applied maximum likelihood models of codon substitution using the program codeml from PAML (*34*, *35*). We found a total of 11 residues (S118, E172, L194, Q198, R245, E339, E391, L392, A416, Q489, R590) to be under positive selection according to the three models tested (table S1). We further used AlphaFold2 to predict the tertiary structure of the $PERU^{DM}$ ectodomain (fig. S11). All residues but one (S118) found to be under positive selection are located on the concave side of the LRR structure, consistent with observations made for binding of the bacterial flagellin epitope flg22 to the *Arabidopsis* LRR-RK FLS2 (*36*). As observed for other immune receptors, diversifying selection may have driven functional diversification of PERU receptors in wild potato populations.

We further observed that Pep-25–insensitive genotypes were found across the American continent ranging from the US to Chile and Argentina (Fig. 4B). By contrast, Pep-25 sensitivity clustered among species belonging to the section Tuberosa or Piurana—which thrive predominantly in the Andean region of Bolivia and Peru (Fig. 4C)—which suggests that the PERU receptor family arose in this region. Wild potatoes carrying $PERU^{LPH}$ alleles also cluster in this region, suggesting that functional diversification of *PERU* alleles in wild potato populations has occurred at its center of origin. *PERU* alleles from multiple potato cultivars used today for crop production all cluster with $PERU^{DM}$ (Fig. 4A and fig. S10A), suggesting that PERU has been maintained during domestication (*13*).

Discussion

In this study, we characterize potato PERU as a bona fide plant PRR conferring *P. infestans* recognition. PERU binds Pep-13/25 patterns that are conserved among species of the genus *Phytophthora*, hetero-dimerizes with BAK1 in a ligand-dependent manner, mediates activation of plant immunity, and increases resistance to a devastating potato disease.

Different ligand response specificities observed among wild *Solanum* accessions indicate that functional diversification within this PRR family has occurred at the site of origin of the predominant allele, $PERU^{DM}$. The explicit use of wild potato populations instead of plant materials that have undergone substantial genetic rearrangements during crop breeding implies that natural forces have been major drivers of immune receptor diversification. The Pep-13 pattern is widespread and highly conserved among plant-associated oomycetes (*10*), a trait that has likely facilitated the evolution of plant PRRs that recognize it. Although residue W231 is invariant in known sequences of *Phytophthora* TGs, polymorphisms affecting the elicitor activity of WT Pep-13 might occur as pathogen pressure in defined ecological niches is assumed to shape immune receptor reservoirs in metazoans and plants (*5*). It is thus conceivable that functional diversification of PERU was driven by escape mutations within *Phytophthora* Pep-13 patterns that enable Pep-13–producing pathogen strains to elude recognition by the predominant allele, $PERU^{DM}$. Microbial evasion strategies to avoid plant immune activation encompass alterations within immunogenic patterns, thus disabling their recognition by plant PRRs (*37*, *38*). In turn, individual plant species have evolved to perceive polymorphic patterns or, alternatively, structurally unrelated immunogenic molecules (*39*–*43*). Likewise, phylogenetically distinct PRRs have evolved in different plant species to recognize structurally unrelated epitopes within individual microbial patterns (*44*, *45*).

We report here the identification of a potato cell surface PRR from the Central Andes and its natural origin in wild potatoes. Our analysis highlights PRR diversification in sympatric, natural potato populations.

REFERENCES AND NOTES

1. B. P. M. Ngou, H.-K. Ahn, P. Ding, J. D. G. Jones, *Nature* **592**, 110–115 (2021).
2. M. Yuan *et al.*, *Nature* **592**, 105–109 (2021).
3. J. Rhodes, C. Zipfel, J. D. G. Jones, B. P. M. Ngou, *Essays Biochem.* **66**, 501–511 (2022).
4. R. N. Pruitt *et al.*, *Nature* **598**, 495–499 (2021).
5. B. P. M. Ngou, R. Heal, M. Wyler, M. W. Schmid, J. D. G. Jones, *Nat. Plants* **8**, 1146–1152 (2022).
6. C. Aguilera-Galvez *et al.*, *Stud. Mycol.* **89**, 105–115 (2018).
7. T. Nürnberger *et al.*, *Cell* **78**, 449–460 (1994).
8. T. Nürnberger, D. Nennstiel, K. Hahlbrock, D. Scheel, *Proc. Natl. Acad. Sci. U.S.A.* **92**, 2338–2342 (1995).
9. V. G. A. A. Vleeshouwers *et al.*, *Annu. Rev. Phytopathol.* **49**, 507–531 (2011).
10. F. Brunner *et al.*, *EMBO J.* **21**, 6681–6688 (2002).
11. B. J. Haas *et al.*, *Nature* **461**, 393–398 (2009).
12. W. E. Fry *et al.*, *Phytopathology* **105**, 966–981 (2015).
13. X. Lin *et al.*, *Front. Plant Sci.* **13**, 1037030 (2023).
14. G. M. Pham *et al.*, *Gigascience* **9**, giaa100 (2020).
15. V. G. A. A. Vleeshouwers *et al.*, *BMC Plant Biol.* **11**, 116 (2011).

16. F. Boutrot, C. Zipfel, *Annu. Rev. Phytopathol.* **55**, 257–286 (2017).
17. W.-L. Wan, K. Fröhlich, R. N. Pruitt, T. Nürnberger, L. Zhang, *Curr. Opin. Plant Biol.* **50**, 18–28 (2019).
18. A. M. Escocard de Azevedo Manhães, F. A. Ortiz-Morea, P. He, L. Shan, *J. Integr. Plant Biol.* **63**, 79–101 (2021).
19. D. Couto, C. Zipfel, *Nat. Rev. Immunol.* **16**, 537–552 (2016).
20. T. A. DeFalco, C. Zipfel, *Mol. Cell* **81**, 3449–3467 (2021).
21. J. Du *et al*., *Nat. Plants* **1**, 15034 (2015).
22. S. H. E. J. Gabriëls *et al*., *Plant J.* **50**, 14–28 (2007).
23. C.-H. Wu, K. Belhaj, T. O. Bozkurt, M. S. Birk, S. Kamoun, *New Phytol.* **209**, 1344–1352 (2016).
24. J. Kourelis *et al*., *PLOS Genet.* **18**, e1010414 (2022).
25. J. Ordon *et al*., *Plant J.* **89**, 155–168 (2017).
26. J. Gantner, J. Ordon, C. Kretschmer, R. Guerois, J. Stuttmann, *Plant Cell* **31**, 2456–2474 (2019).
27. V. G. A. A. Vleeshouwers *et al*., *Eur. J. Plant Pathol.* **105**, 241–250 (1999).
28. S. Lacombe *et al*., *Nat. Biotechnol.* **28**, 365–369 (2010).
29. I. Albert *et al*., *Nat. Plants* **1**, 15140 (2015).
30. D. M. Spooner, M. Ghislain, R. Simon, S. H. Jansky, T. Gavrilenko, *Bot. Rev.* **80**, 283–383 (2014).
31. B. P. M. Ngou, P. Ding, J. D. G. Jones, *Plant Cell* **34**, 1447–1478 (2022).
32. L. Quintana-Murci, A. G. Clark, *Nat. Rev. Immunol.* **13**, 280–293 (2013).
33. P. S. Kahlon, R. Stam, *Curr. Opin. Plant Biol.* **62**, 102040 (2021).
34. Z. Yang, *Mol. Biol. Evol.* **24**, 1586–1591 (2007).
35. Z. Liu *et al*., *Mol. Biol. Evol.* **22**, 659–672 (2005).
36. Y. Sun *et al*., *Science* **342**, 624–628 (2013).
37. P. Buscaill, R. A. L. van der Hoorn, *Plant Cell* **33**, 2116–2130 (2021).
38. Y. Wang, R. N. Pruitt, T. Nürnberger, Y. Wang, *Nat. Rev. Microbiol.* **20**, 449–464 (2022).
39. N. R. Colaianni *et al*., *Cell Host Microbe* **29**, 635–649.e9 (2021).
40. U. Fürst *et al*., *Nat. Plants* **6**, 22–27 (2020).
41. K. Parys *et al*., *Cell Host Microbe* **29**, 620–634.e9 (2021).
42. S. Snoeck *et al*., *eLife* **11**, e81050 (2022).
43. Y. Wei *et al*., *Nat. Commun.* **11**, 3763 (2020).
44. T. Murakami *et al*., *Biosci. Biotechnol. Biochem.* **86**, 490–501 (2022).
45. L. Zhang *et al*., *Nat. Plants* **7**, 1254–1263 (2021).
46. A. Thogani, Git Repository: Phylogenetics of PERU and FLS2 LRR-RK immune receptors in the Solanaceae plant family, version 1, Zenodo (2023); https://doi.org/10.5281/zenodo.8079625.
47. S. Landeo Villanueva, M. C. Malvestiti, W. van Ieperen, M. H. A. J. Joosten, J. A. L. van Kan, *Mol. Plant Pathol.* **22**, 361–372 (2021).
48. D. Tang *et al*., *Nature* **606**, 535–541 (2022).
67. J. Jumper *et al*., *Nature* **596**, 583–589 (2021).

ACKNOWLEDGMENTS

We acknowledge our colleagues T. Bisseling, J. Fliegmann, E. Jacobsen, T. Lahaye, R. Lozano-Durán, J. E. Parker, D. Weigel, and S. Wolf for critical reading. We thank D. Huigen for assistance with crossing and greenhouse work organization; H. Wiegersma and H.Smid for excellent plant care; J. Kourelis and Y. Sugihara for assistance with the phylogenomics, J. Win for advice with PAML analysis; S. Landeo for help with ROS assays; J. Schaart, W. de Martinez and D. Monino Lopez for help with CRISPR-Cas constructs cloning; I. Pereira-Bertram for culturing *Solanum* plants; Matthieu Joosten for VIGS constructs; Geert Kessel for providing *P. infestans* isolates; A. Andriani for help with disease tests; J. Stuttmann for providing *N. benthamiana eds1*, *pad4* and *eds1 pad4* lines, and T. Marti Ferrando for help with molecular work. We recognize the 'Apoplastic interactions' session of the 17th IS-MPMI meeting in Portland 2016 that led to initiating the search for the Pep-13 receptor. We dedicate this work to the late D. Scheel, who started the study on Pep-13. **Funding:** This work was supported by the following: Peruvian Council for science, technology and technological innovation (CONCYTEC) FONDECYT contract 129-2017 (to Y.C.T.A.); Deutsche Forschungsgemeinschaft grant CRC1101-D10 and Nu70/17-1 (to T.N.); The Gatsby Charitable Foundation BBSRC, UK, BBS/E/J/000PR9795; BBSRC, UK, BBS/E/J/000PR9796; BBSRC, UK, BS/E/J/000PR9798.
Author contributions: Conceptualization: Y.C.T.A., A.T., X.L., P.J.W., S.K., V.G.A.A.V., and T.N. Methodology: V.G.A.A.V., T.N., Y.C.T.A., A.T., and A.P. Investigation: Y.C.T.A., L.Z., C.H., A.T., N.J.R., A.P., H.P., and N.S. Resources: A.T., D.W., and M.J.M.P. Visualization: Y.C.T.A., A.T., A.P., and R.B. Writing—original draft: Y.C.T.A., T.N., and V.G.A.A.V. Writing—review and editing: Y.C.T.A., L.Z., A.T., S.K., T.N., and V.G.A.A.V. Funding acquisition: R.G.F.V., S.K., T.N., and V.G.A.A.V. **Competing interests:** S.K. receives funding from industry and cofounded a start-up company (Resurrect Bio Ltd.) on plant disease resistance. S.K. has filed patents on plant disease resistance. All other authors declare that they have no competing interests. **Data and materials availability:** Sequences of cloned homologs were deposited in the National Center for Biotechnology Information database, www.ncbi.nlm.nih.gov (accession nos. OQ999431 to OQ999482). All other data, sequences, datasets, and scripts for the phylogenomic analyses are available in the manuscript, the supplementary material or deposited at Zenodo (*46*). Most of the wild *Solanum* accessions were in possession of WUR Plant Breeding before Nagoya rules applied, part of the material was obtained later from the Center of Genetic Resources, The Netherlands for which we refer to Access and Benefit-Sharing (https://www.absfocalpoint.nl/en/absfocalpoint.htm) at their website (https://cgngenis.wur.nl).
License information: Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.sciencemag.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adg5261
Materials and Methods
Figs. S1 to S11
Tables S1 and S2
References (*49*–*66*)
MDAR Reproducibility Checklist
Data S1 to S10

Submitted 4 January 2023; accepted 11 July 2023
10.1126/science.adg5261

HEART DISEASE

Aging impairs the neurovascular interface in the heart

Julian U. G. Wagner[1,2,3], Lukas S. Tombor[1,2,3], Pedro Felipe Malacarne[4], Lisa-Maria Kettenhausen[1], Josefine Panthel[1], Haris Kujundzic[1], Nivethitha Manickam[1,3], Katja Schmitz[1], Maria Cipca[1], Kathrin A. Stilz[1], Ariane Fischer[1], Marion Muhly-Reinholz[1], Wesley T. Abplanalp[1,2,3], David John[1], Sarajo K. Mohanta[5,6,7], Christian Weber[5,6,7,8], Andreas J. R. Habenicht[5,6,7], Giulia K. Buchmann[4], Stephan Angendohr[9], Ehsan Amin[10], Katharina Scherschel[10,11,12], Nikolaj Klöcker[10], Malte Kelm[9,12], Dominik Schüttler[7,13,14,15], Sebastian Clauss[7,13,14,15], Stefan Günther[3,16], Thomas Boettger[2,3,16], Thomas Braun[2,3,16], Christian Bär[17,18], Minh-Duc Pham[19,20], Jaya Krishnan[1,3,19], Susanne Hille[21,22], Oliver J. Müller[21,22], Tarik Bozoglu[7,23], Christian Kupatt[7,23], Eleonora Nardini[24], Selma Osmanagic-Myers[24], Christian Meyer[11,12], Andreas M. Zeiher[2,3], Ralf P. Brandes[2,3,4], Guillermo Luxán[1,2,3], Stefanie Dimmeler[1,2,3]*

Aging is a major risk factor for impaired cardiovascular health. Because the aging myocardium is characterized by microcirculatory dysfunction, and because nerves align with vessels, we assessed the impact of aging on the cardiac neurovascular interface. We report that aging reduces nerve density in the ventricle and dysregulates vascular-derived neuroregulatory genes. Aging down-regulates microRNA 145 (miR-145) and derepresses the neurorepulsive factor semaphorin-3A. miR-145 deletion, which increased *Sema3a* expression or endothelial *Sema3a* overexpression, reduced axon density, mimicking the aged-heart phenotype. Removal of senescent cells, which accumulated with chronological age in parallel to the decline in nerve density, rescued age-induced denervation, reversed *Sema3a* expression, preserved heart rate patterns, and reduced electrical instability. These data suggest that senescence-mediated regulation of nerve density contributes to age-associated cardiac dysfunction.

The vasculature and nervous system form complex, highly branched networks, which are frequently interdependent and functionally linked. Vessel-nerve alignments are mediated by nerve-derived signals that act on endothelial cells (ECs) or conversely by vascular signals that control nerve-fiber alignment along a preformed vessel template (*1*). Thereby, guidance cues such as semaphorins, Eph/ephrins (erythropoietin-producing hepatocellular carcinoma/Eph receptor-interacting proteins), and vascular endothelial growth factor (VEGF)/VEGF receptor (VEGFR) regulate vessels and neurons and extend or repel axonal growth (*2*). Afferent and efferent cardiac neurotransmission via sympathetic and parasympathetic cardiac nerves modulates many physiological functions of the heart. Thus, imbalances of either branch can lead to arrhythmias. For example, impaired cardiac parasympathetic activity is a negative prognostic indicator and can lead to ventricular arrhythmia (*3*, *4*), whereas both excessive (*5*–*7*) and reduced (*8*) sympathetic activity can also result in arrhythmias. Moreover, cardiac denervation causes silent ischemia and lethal arrhythmia in diabetic hearts (*4*, *8*, *9*). In addition to the control of electrical stability, innervation has other functions and, for example, is essential for regeneration of the heart, as shown in postnatal postinfarction-regeneration models in mice (*10*, *11*). In the vasculature and in other organs, innervation can control inflammation (*12*, *13*). However, whether aging has an impact on cardiac innervation on a cellular and mechanistic level is unknown.

Results

Aging causes declines in sympathetic, parasympathetic, and sensory fibers

We explored the impact of aging on nerve density in old mice. We used 18- to 20-month-old male C57Bl/6J mice, which revealed diastolic dysfunction, although ejection fraction was preserved (fig. S1, A and B). Pan-neuronal staining for class III β-tubulin (TUJ1) showed a robust reduction of axon density in 18-month-old mouse hearts compared with those of 3-month-old male young mice (Fig. 1A). A similar decline in neuron density was observed in old female mice (fig. S2, A and B).

The age-induced reduction in nerve density was specifically detected in the left ventricle (LV), whereas innervation in the right ventricle (RV) was comparable between old and young mice (Fig. 1, B and C). The extent of age-dependent decline in the epicardial region was not as strong as in the subendocardial and myocardial regions, but high-magnification images demonstrated a robust decline of subepicardial axon density in aged hearts as well (Fig. 1B and fig. S3A). This observation was further confirmed in whole-mount staining of old mouse hearts, which showed a decline in nervous fibers across the left wall (Fig. 1D). However, no induction of apoptosis or cell death–related gene expression were detected in old mice stellate ganglia, which are the source of cardiac sympathetic innervation (Fig. S3, B and C, and data S1). To assess the specific time point when denervation starts, we performed a time-course study showing a decline of total nerve density already at 16 months of age with a further decline at 22 months (Fig. 1E and fig. S3, D and E).

Next, we assessed which types of nerves are affected by aging. The heart is innervated by sympathetic, parasympathetic, and sensory fibers, which are commonly stained for tyrosine hydroxylase (TH) (*14*, *15*), choline acetyltransferase (CHAT) (*16*, *17*), and calcitonin gene-related peptide (CGRP) (*18*), respectively. TH-positive nerves were present in both ventricles but were selectively reduced in the LV of aged mice with time kinetics similar to that of total nerves (Fig. 1F and fig. S3, F and G). CHAT-positive nerves were selectively found in the atrioventricular (AV) area and were also reduced in old mouse hearts (Fig. 1G and fig. S3H), whereas innervation was not changed in the sinoatrial node (fig. S3I). The expression of acetylcholinesterase and TH

[1]Institute of Cardiovascular Regeneration, Centre for Molecular Medicine, Goethe University Frankfurt, 60590 Frankfurt, Germany. [2]German Center for Cardiovascular Research (DZHK), Partner Site Rhein-Main, 60590 Frankfurt, Germany. [3]Cardiopulmonary Institute (CPI), 60590 Frankfurt, Germany. [4]Institute for Cardiovascular Physiology, Goethe University Frankfurt, 60590 Frankfurt, Germany. [5]Institute for Cardiovascular Prevention, Ludwig-Maximilians-Universität München (LMU), 80336 Munich, Germany. [6]Munich Cluster for Systems Neurology (SyNergy), 81377 Munich, Germany. [7]German Center for Cardiovascular Research (DZHK), Partner Site Munich, Munich Heart Alliance (MHA), 80802 Munich, Germany. [8]Department of Biochemistry, Cardiovascular Research Institute Maastricht (CARIM), Maastricht University, Maastricht, Netherlands. [9]Department of Cardiology, Pulmonology and Vascular Medicine, Medical Faculty, Heinrich-Heine-University Düsseldorf, 40225 Düsseldorf, Germany. [10]Institute of Neural and Sensory Physiology, Medical Faculty and University Hospital Düsseldorf, Heinrich-Heine-University Düsseldorf, 40225 Düsseldorf, Germany. [11]Division of Cardiology/Angiology/Intensive Care, EVK Düsseldorf, cNEP, cardiac Neuro- and Electrophysiology Research Consortium, 40217 Düsseldorf, Germany. [12]Cardiovascular Research Institute Düsseldorf (CARID), Medical Faculty and University Hospital of Düsseldorf, Heinrich-Heine University Düsseldorf, 40225 Düsseldorf, Germany. [13]Department of Medicine I, University Hospital Munich, Ludwig Maximilian University, 81377 Munich, Germany. [14]Institute of Surgical Research at the Walter-Brendel-Centre of Experimental Medicine, University Hospital, LMU Munich, 81377 Munich, Germany. [15]Interfaculty Center for Endocrine and Cardiovascular Disease Network Modelling and Clinical Transfer (ICON), LMU Munich, 80539 Munich, Germany. [16]Max Planck Institute for Heart and Lung Research, 61231 Bad Nauheim, Germany. [17]Institute of Molecular and Translational Therapeutic Strategies (IMTTS), Hannover Medical School, 30625 Hannover, Germany. [18]REBIRTH-Centre for Translational Regenerative Medicine, Hannover Medical School, 30625 Hannover, Germany. [19]Department of Medicine, Cardiology, Goethe University Hospital, 60590 Frankfurt, Germany. [20]Genome Biologics, 60590 Frankfurt am Main, Germany. [21]Department of Internal Medicine III, University Hospital Schleswig-Holstein, University of Kiel, 24105 Kiel, Germany. [22]German Centre for Cardiovascular Research (partner site Hamburg/Kiel/Lübeck), 24105 Kiel, Germany. [23]Klinik und Poliklinik für Innere Medizin I, University Clinic rechts der Isar, Technical University of Munich, 81675 Munich, Germany. [24]Institute of Medical Chemistry, Center for Pathobiochemistry and Genetics, Medical University of Vienna, A-1090 Vienna, Austria.
*Corresponding author. Email: dimmeler@em.uni-frankfurt.de

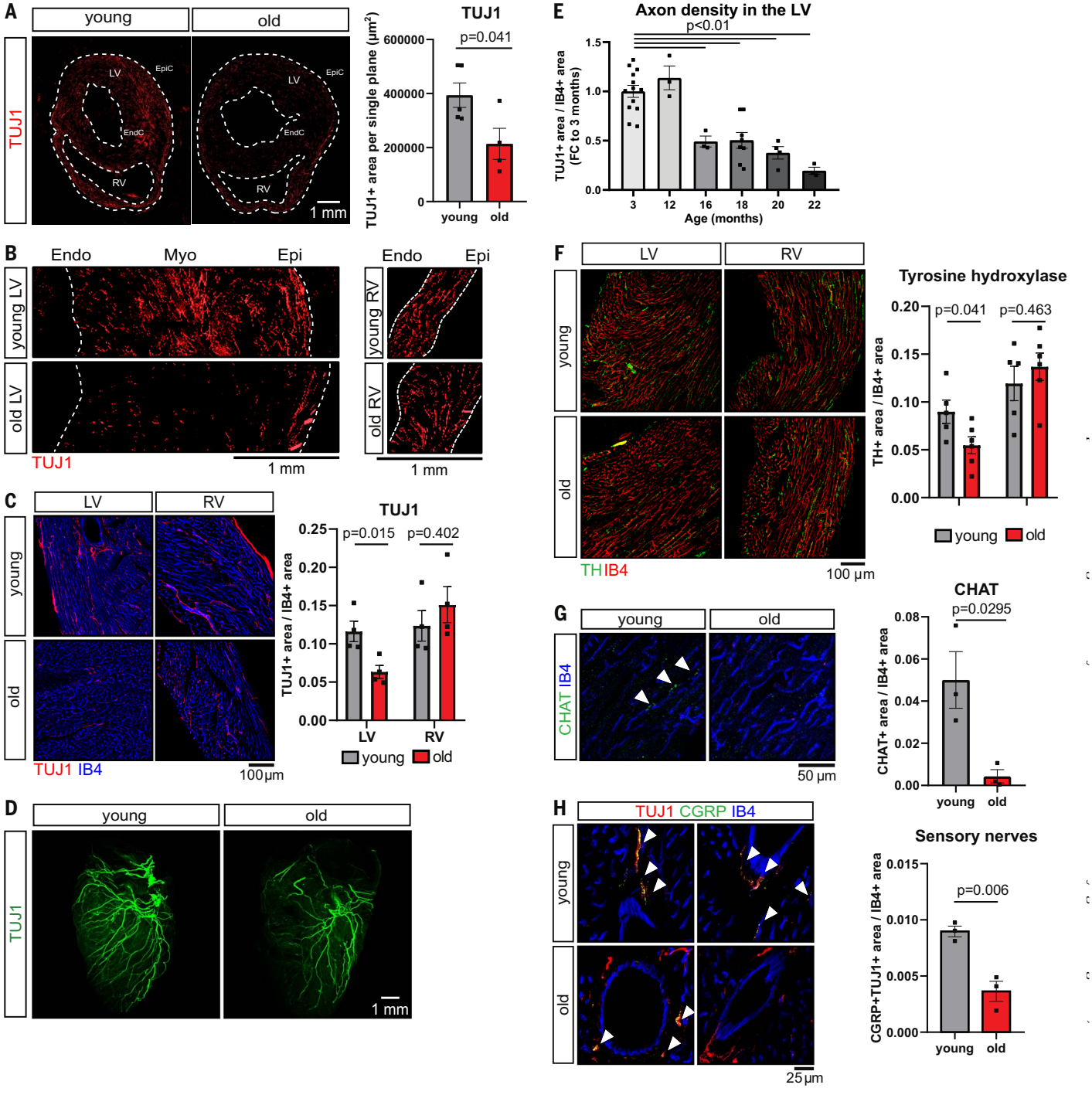


Fig. 1. Left-ventricular axon density is reduced with age. (**A**) Tile scan of 3-month-old (young) versus 18-month-old (old) male mouse heart cross sections [left ventricle (LV), right ventricle (RV)]. Autonomic nervous system is shown by TUJ1 staining (red) and was quantified by TUJ1-positive area per single *z* plane (young, *n* = 5 mice versus old, *n* = 4). (**B**) Zoom-in of the left and right ventricular wall shown in (A). (**C**) Higher magnification of TUJ1-positive area versus IB4-positive area in young (3 months) versus old (20 months) LVs and RVs. Innervation was assessed with TUJ1 (red) and normalized to IB4 (blue) (*n* = 4). (**D**) Whole-mount TUJ1 (green) staining of male young (3 months) and old (18 months) mouse hearts. One representative image per group is shown (*n* = 3). (**E**) Left-ventricular axon density (TUJ1 = green) versus IB4 (red) at 3, 12, 16, 18, 20, and 22 months of age (*n* = 12 versus *n* = 3, *n* = 3, *n* = 8, *n* = 4, and *n* = 3). (**F**) TH staining (green) versus IB4 (red) in male young (3 months) versus old (18 to 20 months) LVs and RVs (young, *n* = 5 versus old, *n* = 6). (**G**) Immunofluorescence (IF) staining for CHAT in young (3 months) versus old (20 months) atrioventricular regions of old mouse hearts. IB4 (red) served as counterstain (*n* = 3). Arrowheads indicate CHAT$^+$ fibers. (**H**) IF staining of sensory nerves [CGRP (green) and TUJ1 (red), indicated by white arrowheads] versus IB4 (blue) in young (3 months) versus old (18 to 20 months) mouse hearts (*n* = 3). Data are shown as means, and error bars indicate the SEM. After passing normality test, *P* values were calculated with the unpaired, two-tailed *t* test [(A), (C), (F), and (H)]. To compare multiple groups, an ordinary one-way analysis of variance (ANOVA) test with a Tukey's post hoc test was used (E).

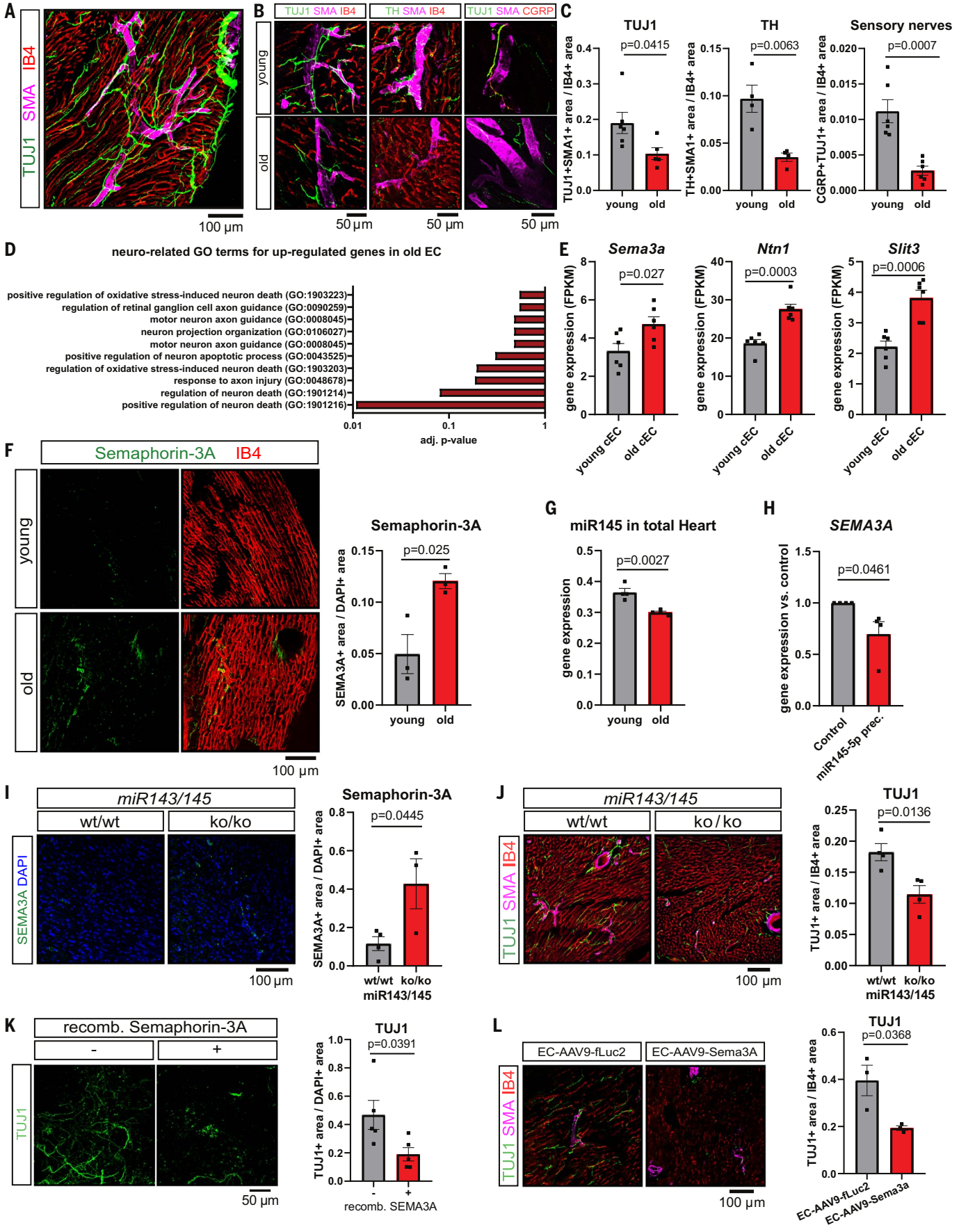

A
TUJ1 SMA IB4
100 µm
B
TUJ1 SMA IB4
TH SMA IB4
TUJ1 SMA CGRP
young
old
50 µm
50 µm
50 µm
C
TUJ1
p=0.0415
TUJ1+SMA1+ area / IB4+ area
young
old
TH
p=0.0063
TH+SMA1+ area / IB4+ area
Sensory nerves
p=0.0007
CGRP+TUJ1+ area / IB4+ area
D
neuro-related GO terms for up-regulated genes in old EC
positive regulation of oxidative stress-induced neuron death (GO:1903223)
regulation of retinal ganglion cell axon guidance (GO:0090259)
motor neuron axon guidance (GO:0008045)
neuron projection organization (GO:0106027)
motor neuron axon guidance (GO:0008045)
positive regulation of neuron apoptotic process (GO:0043525)
regulation of oxidative stress-induced neuron death (GO:1903203)
response to axon injury (GO:0048678)
regulation of neuron death (GO:1901214)
positive regulation of neuron death (GO:1901216)
adj. p-value
E
Sema3a
p=0.027
Ntn1
p=0.0003
Slit3
p=0.0006
gene expression (FPKM)
young cEC
old cEC
F
Semaphorin-3A IB4
young
old
100 µm
Semaphorin-3A
p=0.025
SEMA3A+ area / DAPI+ area
G
miR145 in total Heart
p=0.0027
gene expression
H
SEMA3A
p=0.0461
gene expression vs. control
Control
miR145-5p prec.
I
miR143/145
wt/wt
ko/ko
SEMA3A DAPI
100 µm
Semaphorin-3A
p=0.0445
SEMA3A+ area / DAPI+ area
miR143/145
J
miR143/145
wt/wt
ko/ko
TUJ1 SMA IB4
100 µm
TUJ1
p=0.0136
TUJ1+ area / IB4+ area
K
recomb. Semaphorin-3A
-
+
TUJ1
50 µm
TUJ1
p=0.0391
TUJ1+ area / DAPI+ area
recomb. SEMA3A
L
EC-AAV9-fLuc2
EC-AAV9-Sema3A
TUJ1 SMA IB4
100 µm
TUJ1
p=0.0368
TUJ1+ area / IB4+ area
EC-AAV9-fLuc2
EC-AAV9-Sema3a

Fig. 2. Endothelial-derived semaphorin-3A reduces axon density in the heart. (**A**) Representative image of the neurovascular alignment in a young (3 months) male mouse heart. Nervous fibers were assessed with TUJ1 staining (green), ECs with IB4 (red), and arterioles with smooth muscle actin (SMA, magenta) staining. (**B**) Nervous fibers in the proximal arterial space were assessed in 3-month-old (young) versus 20-month-old (old) mouse hearts with TUJ1 (green) as pan-neuronal marker (n = 6 mice versus n = 5), with TH (green) as sympathetic marker (n = 4 versus n = 4) and CGRP (red) as marker for sensory nerves (n = 6 versus n = 6). (**C**) shows the quantification of the data shown in (B). (**D**) GO analysis of nervous system–related differentially expressed genes in bulk RNA sequencing data of isolated cardiac endothelial cells from young (3 months) versus old (20 months) mice (n = 6 versus n = 6). (**E**) Expression of *Sema3a*, *Ntn1*, and *Slit3* mRNA [fragments per kilobase of transcript per million mapped reads (FPKM)] in cardiac ECs (cECs) isolated from young (3 months) versus old (20 months) mice (n = 6 versus n = 6). (**F**) IF staining of semaphorin-3A (green) and IB4 (red) male in young (3 months) and old (20 months) LVs. (n = 3 versus n = 3). (**G**) Expression of miR-145 in total hearts from young and old mice as assessed with microarray (n = 4 versus n = 4). (**H**) *SEMA3A* mRNA expression in human umbilical-cord vein ECs (HUVECs) upon miR145-5p precursor transfection (n = 4 in vitro experiments versus n = 4). (**I**) IF staining of semaphorin-3A (green) in hearts derived from young (3 months) male and female *miR-143/145*$^{ko/ko}$ mice. SMA (magenta) and IB4 (red) served as control staining (n = 4 mice versus n = 3). (**J**) IF staining of TUJ1 (green) in hearts derived from young (3- to 4-month-old) male and female *miR-143/145*$^{ko/ko}$ mice. SMA (magenta) and IB4 (red) served as control staining (n = 4). (**K**) Human iPS-derived cardiac organoids were treated with human recombinant semaphorin-3A (500 ng/ml) or phosphate-buffered saline as control for 3 days. TUJ1 (green) served as neuronal marker (n = 5 organoids versus n = 5). (**L**) Retargeted adeno-associated virus 9 (AAV9) encoding either fLuc2 (EC-AAV9-fLuc2) or *Sema3a* (EC-AAV9-Sema3a) were intravenously injected to young (3-month-old) male mice. One week after injection, mice were euthanized, and hearts were histologically assessed with TUJ1 (green), SMA (magenta), and IB4 (red) (n = 3 mice versus n = 3). Data are shown as means, and error bars indicate the SEM. After passing normality test, P values were calculated with the unpaired, two-tailed t test.

was not changed in the aging brain (fig. S3J). Sensory nerves were exclusively detected in perivascular regions and were also diminished with age (Fig. 1H). Aging additionally resulted in a higher incidence and severity of ventricular tachycardia in Langendorff-perfused hearts of 18-month-old mice compared with those of young mice (fig. S4), which documented increased electrical instability. These data demonstrate a decline of cardiac innervation in the LV of old-aged mice.

Neurovascular interactions in the aging heart

The vasculature and the nervous system co-develop and remain aligned in the mature heart (Fig. 2A). To address whether the decline in nerve density may be secondary to age-related capillary rarefaction, we histologically assessed capillary density over 22 months (fig. S5A). Capillary density as assessed by isolectin B4 (IB4) or intercellular adhesion molecule 2 showed a decline only at 22 months, whereas capillary perimeter increased (fig. S5). However, nervous fiber density declined at 16 months of age, and vascular alignment of the various nerve types was diminished in 20-month-old mouse hearts (Fig. 2, B and C). These findings suggest that the reduction of nerves is not caused by capillary rarefaction but may be secondary to a dysregulation of capillary function and possibly alterations in neuroguidance cues in the vasculature.

To address whether neuroguidance cues might be dysregulated in the cardiac endothelium of the aging heart, we isolated cardiac ECs from young and old mice and performed RNA sequencing (data S2). Gene Ontology (GO) term analysis of induced genes demonstrated pathways assigned to "neuronal death" and "axon injury" (Fig. 2D and fig. S6A). Genes in these pathways include semaphorin-3A (*Sema3a*) (Fig. 2E), which patterns the autonomic nervous system during development (*19*). Semaphorin-3A is also essential to maintain normal heart rhythm through sympathetic innervation patterning but induces vulnerability to arrhythmias if overexpressed in cardiomyocytes (*5*). In addition, we found up-regulation of members of the Slit/Robo family, such as *Slit3*, which can mediate repulsive signals (*20*), and *Netrin-1* (*Ntn1*), a laminin-related secreted protein (Fig. 2E), which may switch attraction to repulsive responses in a dose-dependent manner (*21*). Recent studies suggest that the combination of guidance factors (such as Slit-family members and Netrin-1) can act in concert to modulate cellular responses (*21*). Validation of protein expression confirmed the up-regulation of semaphorin-3A in aged mouse hearts (Fig. 2F) and also showed that semaphorin-3A is predominantly expressed by vascular cells (fig. S6, B and C).

Endothelial Sema3a is regulated by miR-145 and induces denervation

Because both overexpression and deletion of *Sema3a* may lead to sudden cardiac death and ventricular fibrillation (*6*), we investigated upstream pathways, which may control age-induced induction of *Sema3a*. *Sema3a*-mRNA has two miR-145 binding sites in the 3′ untranslated region (3′UTR) (*22*) (fig. S7A), and miR-145-5p was one of the predicted miRs that was also reduced in the aging heart (Fig. 2G and fig. S7A). Therefore, we hypothesize that miR-145 might repress *Sema3a* in the young heart. Overexpression of miR-145 indeed repressed *SEMA3A* in human umbilical-cord vein ECs in vitro (Fig. 2H). Furthermore, mice in which *miR-143/145* were deleted (*miR-143/145*$^{ko/ko}$ mice) showed increased semaphorin-3A among vessels (Fig. 2I) and reduced axon density (Fig. 2J and fig. S7B) even at a young age (3 to 4 months). To determine the direct effect of semaphorin-3A on neurites, we treated cultured neurons with recombinant semaphorin-3A, which reduced neuronal cell sprouting in vitro (fig. S7C). Recombinant semaphorin-3A further reduced TUJ1^{+} fiber density in human induced pluripotent stem (iPS) cell–derived cardiac organoids (Fig. 2K). Transient endothelial overexpression of *Sema3a* with adeno-associated viral (AAV) vectors, which are coated with an endothelial-targeting peptide linked to polyamidoamine (PAMAM) to enrich for endothelial tropism (*23*), reduced nerve density in vivo (Fig. 2L and fig. S7, D to F). These data suggest that the miR-145-*Sema3a*-axis modulates cardiac nerve density in the aged heart.

Endothelial-cell senescence induces denervation

Sema3a was also up-regulated in senescent ECs, which were generated by continuous passaging to induce replicative senescence as evidenced by acidic β-galactosidase (β-gal) staining (Fig. 3, A and B). Cellular senescence as measured with acidic β-gal staining, CellEvent green detection, and p16^{INK4a} staining, was induced concomitantly in the aging mouse hearts, when neuronal density declines between 16 and 20 months (Fig. 3, C to E, and fig. S8, A and B). Moreover, genetic models of premature senescence, such as fourth-generation *Tert*$^{-/-}$ mice that lack telomerase, confirmed a decline in total and sympathetic nerve density (Fig. 3, F and G, and fig. S8C). By applying a senescence score (*24*) to our previously published single-nuclei RNA (snRNA) sequencing data from young versus old mouse hearts (*25*), we identified ECs as having acquired the most senescent phenotype (fig. S9A). Bulk RNA sequencing data of isolated cardiomyocytes (*26*), fibroblasts (*25*), and ECs confirmed the up-regulation of senescence marker genes predominantly in aged cardiac ECs (fig. S9B). The expression of *Sema3a* and other neurorepelling factors was augmented in ECs with the highest senescence scores (fig. S9C). This suggests that endothelial senescence might contribute to neuronal repulsion or death. Indeed, the selective induction of EC senescence in young animals through endothelial-specific overexpression of progerin (*27*) reduced the density of TUJ1-positive nerves compared with that of control littermates

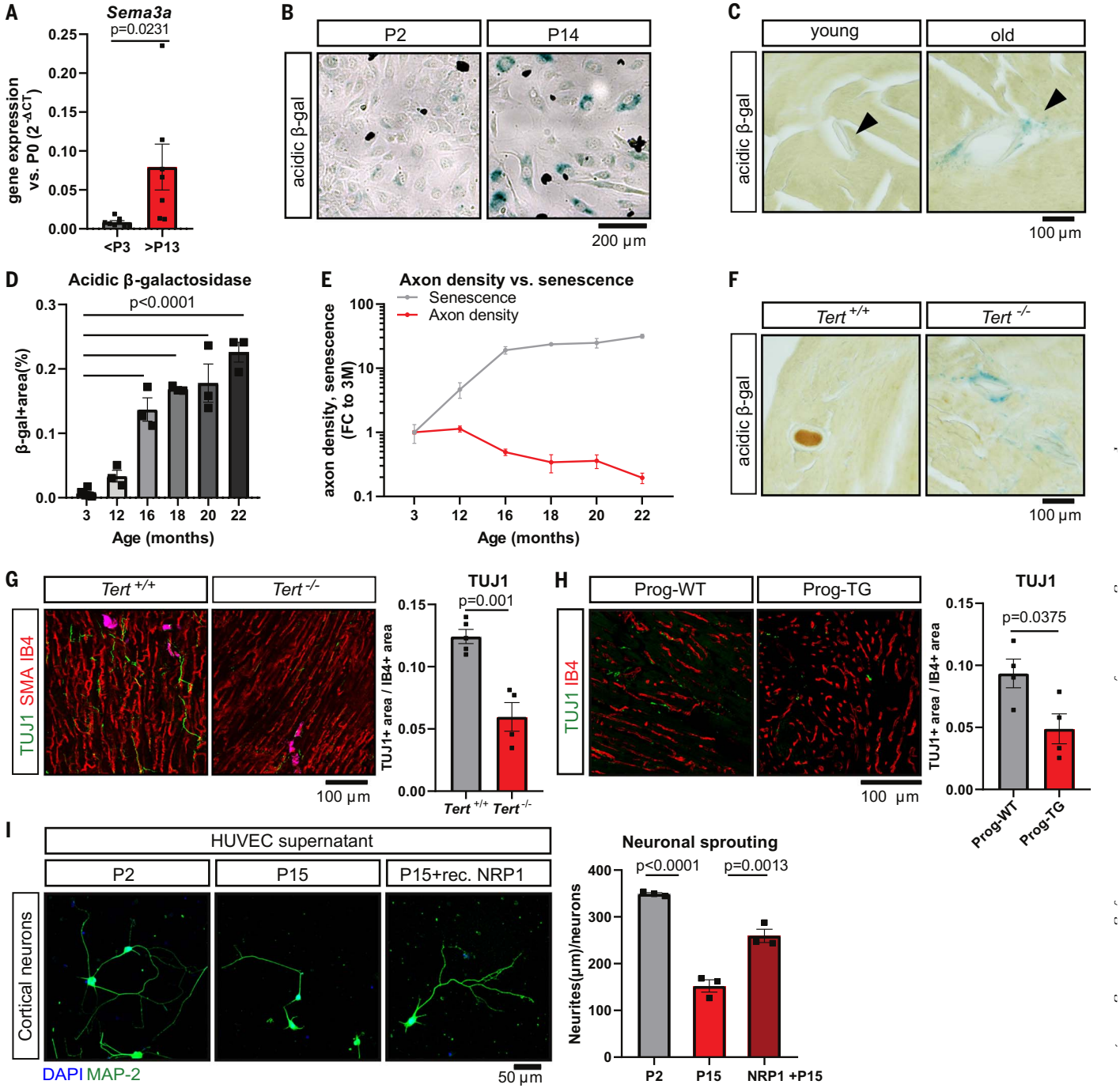


Fig. 3. Senescence impairs cardiac innervation. (**A**) *Sema3a* mRNA expression in short-term passage (P2 to P3) and long-term passage (>P13) HUVECs (n = 8 in vitro experiments versus n = 7). (**B**) Representative image of acidic β-gal staining at P2 and P14 in HUVECs. (**C**) Representative acidic β-gal staining of heart sections from male mice at 3 months and 20 months (3 months, n = 6 mice versus 20 months, n = 3). (**D**) Quantification of acidic β-gal staining of mouse hearts at 3, 12, 16, 18, 20, and 22 months of age (3 months, n = 6 versus 12 months, n = 3; 16 months, n = 3; 18 months, n = 3; 20 months, n = 3; and 22 months, n = 3). (**E**) Axon density (Fig. 1E) and acidic β-gal–positive area (Fig. 3C) plotted in one graph and normalized to 3 months. (**F**) Acidic β-gal staining on heart sections of male *Tert*$^{-/-}$ mice (fourth generation, 3 to 4 months old, n = 4 versus n = 4). (**G**) TUJ1 staining (green) on heart sections of male *Tert*$^{-/-}$ mice (fourth generation, 3 to 4 months old). IB4 (red) and SMA (magenta) served as counterstain (n = 4 versus n = 4). (**H**) TUJ1 staining (green) on heart sections of female and male *Prog* transgenic (*Prog-TG*) mice versus control littermates [*Prog* wild type (Prog-WT) (6 to 7 months)]. IB4 (red) served as counterstain (n = 4 versus n = 4). (**I**) Neurite length analyzed on mouse cortical neurons treated with cell-culture supernatants from HUVECs at P2 or P15. Soluble NRP1 (100 ng/ml) was added to block semaphorin-3A activity. Microtubule-associated protein 2 (MAP-2) was used as neuronal marker (green), and 4′,6-diamidino-2-phenylindole (DAPI) (blue) served as counterstain. Neurite growth was determined by measuring the mean neurite length per neuron (n = 3 in vitro experiments per group). Data are shown as means, and error bars indicate the SEM. After passing normality test, P values were assessed with the unpaired, two-sided t test [(A), (E), (F,) (G), and (H)]. Multiple groups were compared by using a one-way ANOVA with a Tukey's post hoc test [(D) and (I)].

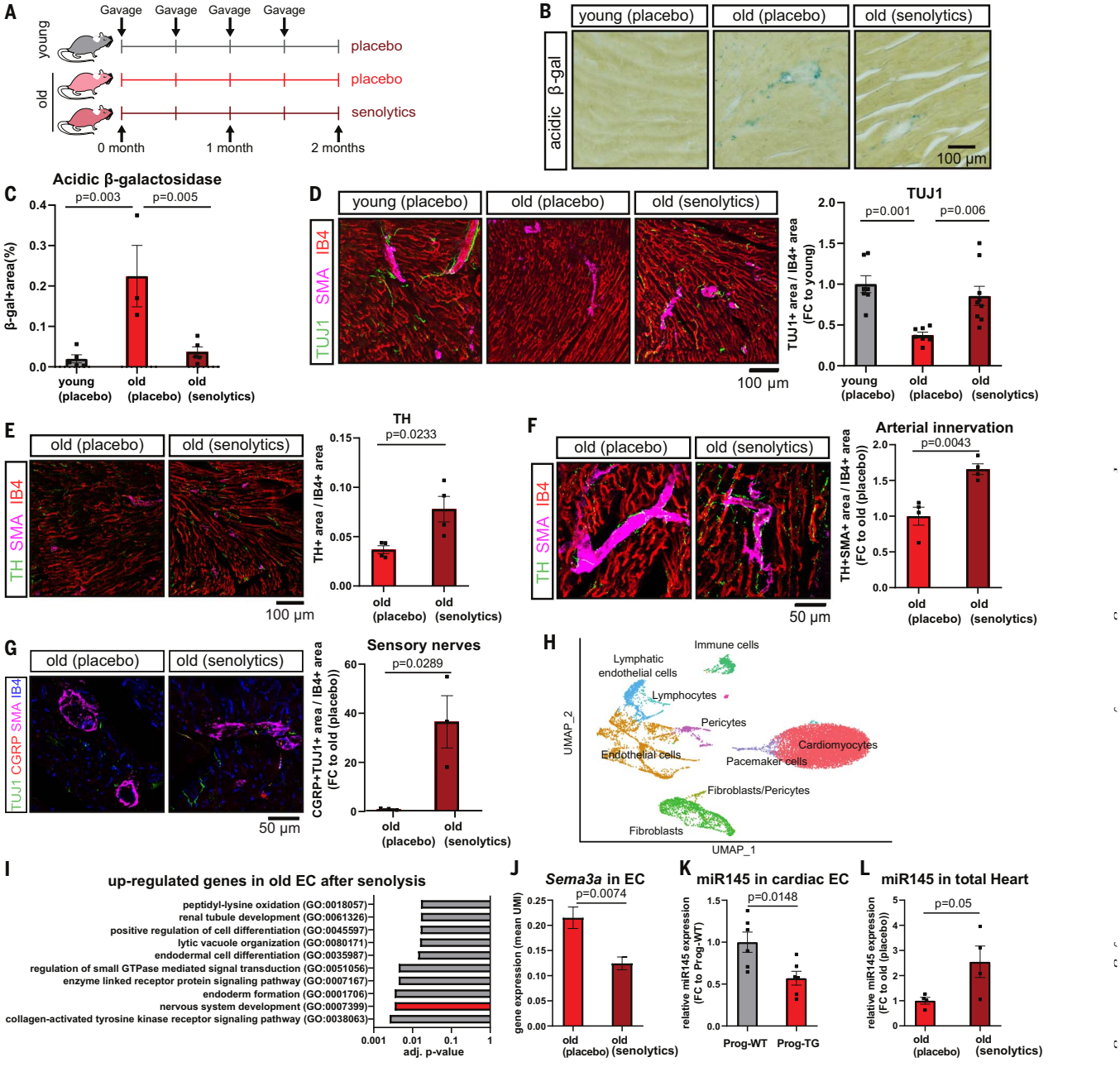


Fig. 4. Senolysis rescues axon density in aged mice. (**A**) Schematic experimental setup. 18- to 19-month-old male mice received a combination of the two senolytic drugs, dasatinib and quercetin, on 3 consecutive days, every second week, for a total duration of 2 months. Young (3 to 4 months) and old mice receiving the vehicle (placebo) served as control cohorts. (**B**) Acidic β-gal staining on heart sections of old male mice after senolytic treatment versus heart sections of the control groups as described in (A). (**C**) The quantification of acidic β-gal positive area [young (placebo), $n = 5$ mice versus old (placebo), $n = 3$ versus old (senolytics), $n = 5$]. (**D**) TUJ1 (green), SMA (magenta), and IB4 (red) staining on heart sections of male old mice after senolytic treatment versus heart sections of the control groups described in (A) [young (placebo), $n = 7$ versus old (placebo), $n = 7$ versus old (senolytics), $n = 9$]. (**E** and **F**) Sympathetic fibers in hearts from old (placebo) versus old (senolytics) mice by TH staining (green). IB4 (red) and SMA (magenta) served as counterstain. (E) shows the overall sympathetic fiber density, and (F) shows the arterial innervation ($n = 4$ versus $n = 4$). (**G**) Sensory nerves in the arterial space were assessed with TUJ1 (green) and CGRP (red). SMA (magenta) and IB4 (blue) served as counterstain ($n = 3$). (**H**) Uniform manifold approximation and projection (UMAP) visualizing snRNA sequencing of old male mice after placebo and senolytic treatment ($n = 3$ versus $n = 3$). (**I**) GO terms analysis of significantly up-regulated genes (cutoff, $P < 0.05$) in ECs upon senolytic treatment. Listed are the top 10 regulated GO terms. (**J**) *Sema3a* mRNA [mean unique molecular identifier (UMI)] expression in the endothelial cluster of the snRNA sequencing data shown in (H). (**K**) Total RNA was extracted from isolated cardiac ECs from *Prog-WT* versus *Prog-TG* mice. Expression of miR-145 was determined with TaqMan polymerase chain reaction (PCR). U6 served as a housekeeping gene ($n = 6$ versus $n = 6$). (**L**) Total RNA was isolated from whole-heart samples of placebo versus senolytics-treated old mice. Expression of miR-145 was determined with TaqMan PCR. U6 served as a housekeeping gene ($n = 4$ versus $n = 4$). Data are shown as mean, and error bars indicate the SEM. After passing normality test, P values were assessed with the unpaired, two-tailed t test [(E), (F), (G), (K), and (L)]. For comparing multiple groups, an ordinary one-way ANOVA with a Tukey's post hoc Tukey test was used [(B) to (D)]. To assess the P value of snRNA sequencing data, a cluster t test was used (J).

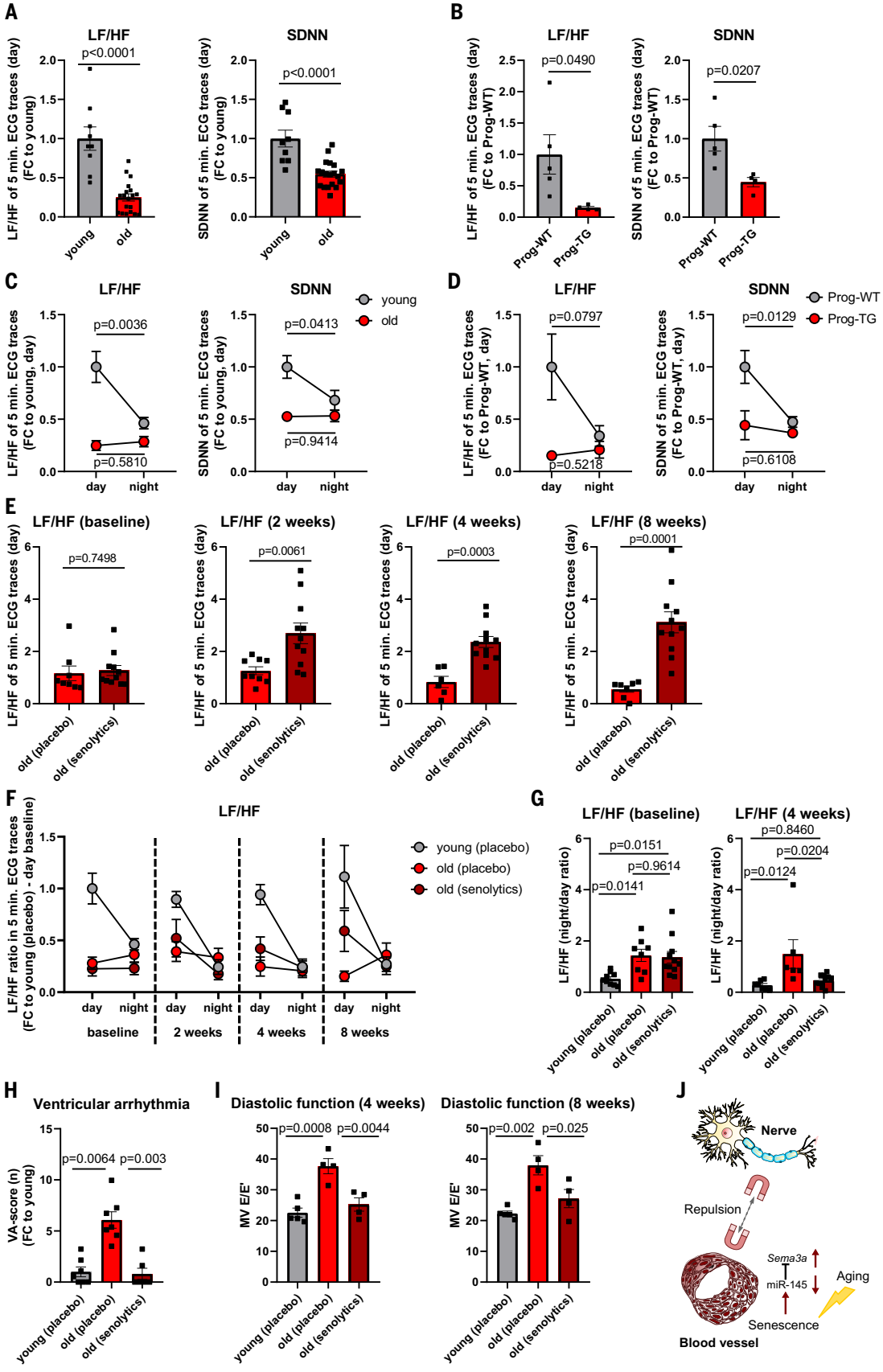


Fig. 5. Heart rate variability (HRV) is altered by senescence. (**A**) HRV was assessed as LF/HF ratio and SDNN remotely in 3-month-old versus 18-month-old male mice by telemetry (n = 9 versus n = 19). (**B**) HRV was assessed as LF/HF ratio and SDNN remotely in 3- to 4-month-old male and female *Prog-TG* mice versus control littermates (*Prog-WT*) (n = 5 versus n = 4). (**C** and **D**) HRV (LF/HF and SDNN) assessed from young versus old mice [(C), n = 9 versus n = 19] and from *Prog-WT* versus *Prog-TG* mice [(D), n = 5 versus n = 4] at day and night. (**E** to **G**) 18-month-old mice were treated with 5 mg/kg dasatinib and 50 mg/kg quercetin (senolytics) on 3 consecutive days, every second week for 2 months. (E) HRV (LF/HF) was assessed remotely with telemetry. Three-month-old (young) and 18-month-old (old) mice receiving the solvent as placebo served as control groups (n = 6 to 9 versus n = 8 versus n = 11). LF/HF was assessed at baseline and at 2, 4, and 8 weeks of senolytic treatment. (F) LF/HF ratio during day and night. (G) Night/day ratio of LF/HF at baseline and 4 weeks upon senolytic treatment. (**H**) Ventricular arrhythmia (VA) score ex vivo in hearts from old mice upon senolysis. Young and old mice receiving placebo served as control groups (n = 7, n = 7, and n = 6). (**I**) Diastolic function was assessed as mitral valve E/E' [E-wave velocity (E, early diastolic filling velocity)/early diastolic tissue velocity (E')] in male young (3 months) versus old (18 months) mice after 4 and 8 weeks of placebo and senolytic treatment (n = 5 versus n = 4 versus n = 4). (**J**) Graphical summary. Data are shown as mean, and error bars indicate the SEM. After passing normality test, P values were assessed-with the unpaired, two-tailed t test [(A) to (E)]. For multiple group comparison, an ordinary one-way ANOVA with a Tukey's post hoc test was used [(G) and (H)]. Data shown in (H) are not normally distributed (Kolmogorov-Smirnov or Shapiro-Wilk normality test), and P values were determined with a Kruskal-Wallis test with a Dunn's multiple comparison test.

(Fig. 3H). Different models of premature senescence indicate that the induction of (endothelial) senescence is sufficient to induce cardiac sympathetic denervation.

To determine whether senescent EC-derived semaphorin-3A directly affects nerves, we tested the effects of supernatants of long-term cultured senescent human umbilical venous ECs or of cardiac ECs derived from young versus old mice on neurons. The supernatants of senescent ECs reduced neuronal sprouting (Fig. 3I and fig. S10A). This effect was partially prevented by blocking semaphorin-3A with soluble recombinant neuropilin-1 (NRP1), which acts as a decoy to inhibit semaphorin-3A activity (*28*) and by using anti-semaphorin-3A neutralizing antibodies (Fig. 3I and Fig. S10B), which confirmed that senescent endothelial-derived semaphorin-3A can indeed disturb neurite guidance in vitro.

Senolytic treatment reduces the age-associated denervation

To determine whether interfering with cellular senescence in vivo might prevent cardiac denervation in the aged heart, we treated old mice with 5 mg/kg dasatinib and 50 mg/kg quercetin, a combination of senolytics that reduces the number of senescent cells by targeting anti-apoptotic pathways and expands life span in vivo (*29*, *30*). Treatment was applied by means of oral gavage to aged mice (18 months) on 3 consecutive days, every second week for a total duration of 2 months (Fig. 4A). At 2 months after the start of treatment, the number of senescent acidic β-gal–positive and $p16^{INK4a}$-positive cells were lower as compared with those in placebo-treated controls (Fig. 4, B and C, and fig. S11A). The reduction in senescent cells from senolytic treatment was paralleled by a rescue of total nerve density as well as sympathetic and sensory nerves in the hearts of old mice (Fig. 4, D to G). Senolytic treatment additionally abrogated the age-associated impairment of endothelial function because it inhibited age-induced increases in capillary perimeter and improved outgrowth of aortic ECs ex vivo (fig. S11, B and C).

To provide mechanistic insights into how senolytics rescue cardiac innervation, we performed snRNA sequencing of old mouse hearts treated with senolytics or placebo (Fig. 4H). Senolytic treatment affected genes associated with "nervous system development" within the top-regulated genes of cardiac ECs (Fig. 4I and data S3). Moreover, *Sema3a*, which we showed to be derepressed in the aging heart, was reduced in old-heart ECs but not in non-ECs such as fibroblasts after senolytic treatment (Fig. 4J and fig. S12). The expression of miR-145, which is reduced in aging hearts (Fig. 2G) and in endothelial progeria mice (Fig. 4K), was partially rescued upon senolysis (Fig. 4L). These data suggest that the reduction of senescent cells in aged hearts augments miR-145 expression and reduces *Sema3a* expression, which potentially restores cardiac innervation.

To determine the effect of aging, endothelial senescence, and senolytic treatment on autonomic and electrical instability, we additionally measured heart rate variability (HRV), which is indicative of an impairment of both parasympathetic and sympathetic innervation in the elderly and has a negative prognostic value (*31*). Increased age, induction of endothelial senescence, or endothelial overexpression of *Sema3a* in 3- to 4-month-old mice reduced HRV as assessed by the ratio of low frequency to high frequency (LF/HF) and standard deviation of normal-to-normal (SDNN) heartbeat intervals (Fig. 5, A and B, and fig. S13, A and B). All three conditions also impaired the day-night rhythm (Fig. 5, C and D, and fig. S13, A and B). Senolytic treatment reversed the age-associated decline in HRV as assessed from the LF/HF ratio and SDNN as early as 2 weeks after the start of treatment (Fig. 5E and fig. S13C). Two months of senolytic treatment further improved the LF/HF ratio in aged mice and restored the characteristic day-night rhythm, whereas old control mice further deteriorated in their autonomous function (Fig. 5, F and G). Additionally, senolytics-treated old mice were slightly more responsive to sympathetic activation by atropine and showed a higher response to the blockade of autonomic control by the combination of atropine and metoprolol, even though the physical activity of these mice was not significantly changed (fig. S13, D to F). In addition, senolytic treatment reduced vulnerability to arrhythmia 8 weeks after treatment (Fig. 5H). These data indicate that senolytic treatment induces a reinnervation of the aging heart, restores the sympathovagal balance, and reduces electrical instability.

Because aging is associated with a decline in diastolic function, we further addressed the effects of senolytics on cardiac function by using echocardiography. Indeed, senolytic treatment improved cardiac function as evidenced by a normalized diastolic function at 4 and 8 weeks of treatment (Fig. 5I, fig. S14, and table S1). Senolytic treatment additionally reduced cardiac fibrosis (fig. S15A). To address whether denervation and senescence may be linked to cardiac fibrosis, we determined cardiac fibrosis in our other models. However, neither miR-143/miR-145 deficiency nor premature senescence induced fibrosis (fig. S15, B to D). By contrast, pharmacological short-term denervation by 6-hydroxydopamine (6-OHDA) augmented cardiac fibrosis (fig. S16). Furthermore, we determined the effects of senolytics on cardiac inflammation. Whereas senolytic treatment reduced the age-associated increase in macrophage accumulation (fig. S17A), miR-143/miR-145 deficiency or endothelial overexpression of *Sema3a* did not affect macrophage numbers (fig. S17, B and C).

Discussion

In this study, we demonstrated that aging reduces axon density in the heart. Aging-induced decline in axon density was associated with reduced miR-145 and derepression of its target, the neuronal repulsive signal semaphorin-3A, which induces electrical instability in the heart (*5*) (Fig. 5J). Unexpectedly, induction of cellular senescence, which is a hallmark of aging, was inversely correlated with the onset of axon decline. Targeting senescent cells pharmacologically was sufficient to prevent the decline in axon density and reduced *Sema3a* expression in the aging heart, suggesting a key role of senescent cells in cardiac denervation. Senescent cells release numerous secreted factors, termed "senescence-associated secretory phenotype" (SASP), which profoundly alter the microenvironment in the aging heart (*32*, *33*). Although neuronal guidance factors have not been reported as part of the general SASP, semaphorin-3A is induced in senescent ECs and may represent a specific component of vascular SASP in the aging heart.

The question of which cell type(s) further contribute(s) to the observed effects will need further studies. We demonstrated that the selective induction of premature aging in ECs is sufficient to reduce nerve density. However, we cannot exclude the involvement of other cells in this process, especially because senolytic drugs may affect a broad range of cells. Although *Sema3a* was lowly expressed and not significantly regulated by senolytics in nonECs, neuronal- and axonal-related pathways were found within the top 25 regulated GO terms in some fibroblast clusters of mice treated with senolytics. These findings indicate that ECs may play a critical role in age-related denervation, but other cells such as pericytes or fibroblasts may also contribute to the observed phenotype.

Our study additionally demonstrates that senolytic treatment reduces vulnerability to arrhythmia and normalizes HRV and the circadian rhythm. A decline in HRV is typically observed in the elderly, is indicative of impaired sympathetic and parasympathetic innervation, and is associated with increased electrical instability leading to increased overall mortality (*31*). Our finding that senolytic treatment normalizes HRV during aging therefore supports a functional benefit of the treatment. HRV and vulnerability to arrhythmias are controlled by various additional mechanisms, which might also be affected by senolytics. Thus, our study does not claim that the effects of senolytics on the electrical control of the heart are entirely mediated by the restoration of cardiac innervation. Additionally, there is no doubt that aging also affects autonomic dysfunction in the brain.

However, the demonstration of a decrease in HRV in endothelial progeria mice supports the concept that EC senescence can affect the autonomous control of the heart.

Innervation not only is important for the control of heart rhythm but also provides critical paracrine factors, which, for example, contribute to cardiac regeneration (*10*, *11*). A decline in nerve density may consequently lead to depletion of nerve-derived factors influencing the reparative function of the heart, as was demonstrated for myocardial infarction in adult mice (*34*). Moreover, in other tissues, nerves interact with immune cells (*35*) and can control vascular inflammation (*12*). Senolytic treatment reduced the age-associated accumulation of macrophages and fibrosis in the heart. However, miR-143/miR-145 deficiency or endothelial overexpression of *Sema3a* did not affect macrophage numbers in the heart at a time point when the decline in nerves was already profound. Additionally, neither miR-143/miR-145 deficiency nor premature senescence induced cardiac fibrosis. These data suggest that cardiac inflammation or tissue fibrosis is unlikely the direct cause of denervation. However, the rescue of denervation and an anti-inflammatory and antifibrotic response to senolytic treatment might together contribute to its overall beneficial anti-aging effects. Short-term pharmacological denervation by 6-OHDA indeed augments cardiac fibrosis. The specific contribution of the sympathetic nervous system to the overall beneficial effects of senolytics deserves further in-depth analysis—for example, whether denervation over a prolonged period may affect inflammation and contribute to "inflammaging" and cardiac dysfunction. The findings we have presented may lay the ground to decipher neuronal cross-talk in the heart and its role in aging.

REFERENCES AND NOTES

1. B. Peguera, M. Segarra, A. Acker-Palmer, *Curr. Opin. Neurobiol.* **69**, 202–213 (2021).
2. A. Eichmann, F. Le Noble, M. Autiero, P. Carmeliet, *Curr. Opin. Neurobiol.* **15**, 108–115 (2005).
3. M. J. Shen, D. P. Zipes, *Circ. Res.* **114**, 1004–1021 (2014).
4. C. Jungen *et al*., *Nat. Commun.* **8**, 14155 (2017).
5. M. Ieda *et al*., *Nat. Med.* **13**, 604–612 (2007).
6. K. Fukuda, H. Kanazawa, Y. Aizawa, J. L. Ardell, K. Shivkumar, *Circ. Res.* **116**, 2005–2019 (2015).
7. P. J. Schwartz, *Nat. Rev. Cardiol.* **11**, 346–353 (2014).
8. C. Jungen *et al*., *Am. J. Physiol. Heart Circ. Physiol.* **317**, H1328–H1341 (2019).
9. M. Ieda *et al*., *Circulation* **114**, 2351–2363 (2006).
10. A. I. Mahmoud *et al*., *Dev. Cell* **34**, 387–399 (2015).
11. I. A. White, J. Gordon, W. Balkan, J. M. Hare, *Circ. Res.* **117**, 990–994 (2015).
12. S. K. Mohanta *et al*., *Nature* **605**, 152–159 (2022).
13. D. Brea, H. Veiga-Fernandes, *Nature* **602**, 217–218 (2022).
14. B. Zhang *et al*., *Nature* **577**, 676–681 (2020).
15. F. Cardoso *et al*., *Nature* **597**, 410–414 (2021).
16. A. Veerakumar, A. R. Yung, Y. Liu, M. A. Krasnow, *Nature* **606**, 739–746 (2022).
17. P. S. Rajendran *et al*., *Nat. Commun.* **10**, 1944 (2019).
18. D. B. Mach *et al*., *Neuroscience* **113**, 155–166 (2002).
19. D. Carulli, F. de Winter, J. Verhaagen, *Front. Synaptic Neurosci.* **13**, 672891 (2021).
20. K. Brose *et al*., *Cell* **96**, 795–806 (1999).
21. I. Dupin, L. Lokmane, M. Dahan, S. Garel, V. Studer, *Neural Dev.* **10**, 5 (2015).
22. M. Rezaeepoor *et al*., *Cell J.* **20**, 113–119 (2018).
23. T. Bozoglu *et al*., *Adv. Sci.* **9**, 2103867 (2022).
24. T. Kiss *et al*., *Geroscience* **42**, 429–444 (2020).
25. R. Vidal *et al*., *JCI Insight* **4**, e131092 (2019).
26. D. J. Trembinski *et al*., *Nat. Commun.* **11**, 2039 (2020).
27. C. Manakanatas *et al*., *Aging* **14**, 195–224 (2022).
28. T. Kawasaki *et al*., *Development* **129**, 671–680 (2002).
29. W. A. Owens, A. Walaszczyk, I. Spyridopoulos, E. Dookun, G. D. Richardson, *Mech. Ageing Dev.* **198**, 111540 (2021).
30. Y. Zhu *et al*., *Aging Cell* **14**, 644–658 (2015).
31. P. Nicolini, M. M. Ciulla, C. De Asmundis, F. Magrini, P. Brugada, *Pacing Clin. Electrophysiol.* **35**, 621–638 (2012).
32. M. Mehdizadeh, M. Aguilar, E. Thorin, G. Ferbeyre, S. Nattel, *Nat. Rev. Cardiol.* **19**, 250–264 (2022).
33. E. Tzahor, S. Dimmeler, *Science* **377**, eabm4443 (2022).
34. K. A. Ziegler *et al*., *Cardiovasc. Res.* **114**, 291–299 (2018).
35. C. Godinho-Silva, F. Cardoso, H. Veiga-Fernandes, *Annu. Rev. Immunol.* **37**, 19–46 (2019).

ACKNOWLEDGMENTS

We thank A. Wilken-Schmitz, F. Vetter, and T. Sun for experimental support. **Funding**: The study is supported by the European Research Council (ERC-2021-ADG, GAP-101053352, Neuroheart to S.D.), the German Research Foundation (SFB1366, Project B4; and TRR267, Project B3, the Cluster of Excellence Cardiopulmonary Institute Exc2026/1 to S.D.; SFB1123, Project Z1 to S.K.M.; and the Munich Cluster for Systems Neurology, Project C6 to C.W.), the German Center for Cardiovascular Research [DZHK Shared Expertise (B22-014 SE) to S.C.], and the Dr. Rolf M. Schwiete Foundation (2021-002 to J.W.). Views and opinions expressed are those of the authors only and do not necessarily reflect those of the funding agencies. **Author contributions:** J.U.G.W. and S.D. planned the project and wrote the manuscript. J.U.G.W., L.M.K., and J.P. conducted the majority of experiments. S.G. performed bulk RNA sequencing. L.S.T., W.T.A., and D.J. performed single-nuclei analysis. W.T.A. performed snRNA sequencing. L.M.K., J.P., M.C., E.N., and K.A.S. contributed to transcriptomic analyses. N.M. performed neuron culture experiments. A.F., P.F.M., K.S., and R.P.B. contributed to telemetry studies. L.M.K., M.C., H.K., G.K.B., M.M.R., and G.L. contributed to histology and whole-mount staining. S.C., D.S., S.A., E.A., N.K., M.K., and C.M. contributed to electrophysiology and telemetry studies. T.Boe. and T.Br. provided miR-145 mice. C.B. provided samples of *Tert* mice. E.N. and S.O.M. provided *Prog-Tg* mice and performed preassessment analysis. M.D.P. and J.K. provided cardiac organoids. S.H., O.J.M., T.Boz., and C.K. provided AVV9. G.L., C.M., T.Br., R.P.B., and A.M.Z. provided conceptual advice. S.K.M., A.J.R.H., and C.W. performed OHDA denervation and provided denervated and control heart samples. **Competing interests:** C.K. and T.Boz. are inventors on a patent application (PCT/EP2022/086634) submitted by the authors that covers the endothelial targeting vector. All other authors declare that they have no competing interests. **Data and materials availability:** All data are available in the manuscript or the supplementary materials. The single-cell experiment is available at ArrayExpress under the accession no. E-MATB13093. Material is available upon request from the authors indicated in the supplementary material. The endothelial targeting vector is available from C.K. under a materials transfer agreement with Klinikum rechts der Isar der Technischen Universität München. **License information:** Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.ade4961
Materials and Methods
Figs. S1 to S18
Table S1
References (*36*–*43*)
Data S1 to S3
MDAR Reproducibility Checklist

Submitted 20 August 2022; resubmitted 11 April 2023
Accepted 11 July 2023
10.1126/science.ade4961

SENSATION

PIEZO2 and perineal mechanosensation are essential for sexual function

Ruby M. Lam[1,2], Lars J. von Buchholtz[3], Melanie Falgairolle[1], Jennifer Osborne[1], Eleni Frangos[1], M. Rocio Servin-Vences[4], Maximilian Nagel[1], Minh Q. Nguyen[3], Monessha Jayabalan[1], Dimah Saade[5], Ardem Patapoutian[4], Carsten G. Bönnemann[5], Nicholas J. P. Ryba[3], Alexander T. Chesler[1,5]*

Despite the potential importance of genital mechanosensation for sexual reproduction, little is known about how perineal touch influences mating. We explored how mechanosensation affords exquisite awareness of the genitals and controls reproduction in mice and humans. Using genetic strategies and in vivo functional imaging, we demonstrated that the mechanosensitive ion channel PIEZO2 (piezo-type mechanosensitive ion channel component 2) is necessary for behavioral sensitivity to perineal touch. PIEZO2 function is needed for triggering a touch-evoked erection reflex and successful mating in both male and female mice. Humans with complete loss of PIEZO2 function have genital hyposensitivity and experience no direct pleasure from gentle touch or vibration. Together, our results help explain how perineal mechanoreceptors detect the gentlest of stimuli and trigger physiologically important sexual responses, thus providing a platform for exploring the sensory basis of sexual pleasure and its relationship to affective touch.

Sexual reproduction is a fundamental driver for animal behavior, and adaptations required for courtship, including sexual ornamentation and ritual displays, are cornerstones of evolutionary theory (*1*–*3*). Visual, auditory, and olfactory cues promote mating in various mammalian species (*4*–*8*); however, the act of copulation itself can be considered a specialized form of touch endowed with its own cortical field (*9*). Although the discovery of the mechanically gated ion-channel PIEZO2 (piezo-type mechanosensitive ion channel component 2) (*10*) has spurred remarkable progress in our understanding of discriminative touch (*11*, *12*), far less is known about mechanosensation in the genitals (*13*, *14*), including how it triggers physiological responses and elicits pleasure. We hypothesized that sexual touch might exhibit unusual response specialization to control mating and provide affective and motivational feedback. We also expected that there would be sexual dimorphism both in sensation and in responses triggered by genital-innervating mechanosensors. To test these hypotheses, we developed a series of behavioral and functional imaging assays to probe the role of PIEZO2 in genital mechanosensation and sexual function. In addition, by exploring the impact of PIEZO2 loss of function caused by a rare inherited syndrome, we determined how these findings relate to human sexual experience.

[1]National Center for Complementary and Integrative Health (NCCIH), Bethesda, MD 20892, USA. [2]Brown-National Institutes of Health Graduate Partnerships Program, Brown University, Providence, RI 02912, USA. [3]National Institute of Dental and Craniofacial Research, Bethesda, MD 20892, USA. [4]Howard Hughes Medical Institute, Department of Neuroscience, Dorris Neuroscience Center, The Scripps Research Institute, La Jolla, CA 92037, USA. [5]National Institute of Neurological Disorders and Stroke, Bethesda, MD 20892, USA.
*Corresponding author. Email: alexander.chesler@nih.gov

Unusual sensitivity and PIEZO2 dependence of perineal touch

A standard touch sensitivity test uses calibrated von Frey filaments to measure detection threshold. In mice, von Frey sensitivity of the glabrous hind-paw and hairy skin of the face have similar withdrawal thresholds (*15*–*17*) despite very different patterns of innervation (*18*). We adapted this assay to compare stimulation of the perineum (the region extending from the anus to the genitals in male and female mice) with that of the plantar surface of the paw. In the hind-paw assay, mice respond by withdrawing the paw with no indication of pain or distress. Our data (Fig. 1A) match literature reports, with filaments ≥0.4 g eliciting responses in the majority of trials, but filaments ≤0.16 g rarely provoking reaction (*15*, *17*). By contrast, stimulation of the perineum evoked a highly stereotyped startle and investigative response (movie S1) both in male and female mice. Even the finest filament available (0.008 g) elicited this reaction from every animal (Fig. 1B), demonstrating exquisite sensitivity of the perineum to forces below those that reliably trigger responses from other sites, even in mice with profound allodynia (*15*, *17*); female mice were marginally but consistently more sensitive than males (Fig. 1B).

The mechanically activated ion channel PIEZO2 is essential for discriminative touch in mice and humans (*11*, *12*). We anticipated that this mechanoreceptor would also be responsible for the sensitivity of the perineum. *Piezo2*-null mice die as neonates (*19*); therefore, we generated conditional genetic deletions using a *Hoxb8-Cre* line (Piezo2^{Hoxb8}) to target cells below the mid-thoracic region (*17*). We used this strategy to assess the role of PIEZO2 in perigenital sensation and observed profound loss of behavioral response to von Frey filaments (Fig. 1C), with the highest force tested (1.4 g) eliciting responses in only ~40% of trials (movie S1). Local inhibition of the perineum with lidocaine attenuated von Frey responses of controls (fig. S1), and Piezo2^{Hoxb8} responses to noxious mechanical pinprick were indistinguishable from those of controls (Fig. 1C and movie S1). Therefore, the Piezo2^{Hoxb8} deficit is likely to be sensory rather than related to a movement disorder (*20*). These experiments demonstrated that PIEZO2 is crucial for triggering behavioral responses to the gentlest of perigenital touch in mice; without this touch receptor, von Frey stimulation of the genital region rarely elicited responses even at intensities considered noxious.

We previously studied a rare cohort of people with biallelic loss-of-function variants of *PIEZO2* who have sensory deficits fully consistent with those described in animal models (*11*, *21*). In our clinical interviews, five adult human subjects with PIEZO2-deficiency syndrome (three male and two female) reported severe hyposensitivity in genital sensation (table S1); however, comprehensive quantitative testing has not been possible. One individual adult male consented to quantitative sensory testing of his genitalia during clinical evaluation. His penile von Frey detection threshold (3.1 ± 1.5 g) was far higher than values reported in the literature: 0.3 to 0.6 g in a similar location (*22*). He had difficulty detecting pressure below 1 kg/cm^2 at the midshaft and was insensitive to strong vibration at 50 and 100 Hz, which is consistent with our findings in mice. By contrast, literature values for penile fine-touch pressure thresholds in a range of healthy men are far lower (*23*), and vibration is normally readily detected (*23*).

Anatomy of perineal neurons

Somatosensory neurons in the lower body have soma in lumbar (L1 to L6) and sacral (S1 to S4) dorsal root ganglia (DRG) (*24*). However, few details about the types or sensitivity of neurons that target the genitals are known. Multicolor cholera toxin subunit-b (CTB) tracing from both hind-paw and genitals robustly labeled neurons in S1 and S2 DRG (fig. S2A) and distinguished neurons that target perigenital subregions (Fig. 1, D and E). Injections to the perineum, prepuce, and glans (male mice) or vaginal opening (female mice) resulted in largely nonoverlapping labeling of neurons with a range of cell diameters (Fig. 1D and fig. S2B). In both sexes, dense projections targeted the (L6 to S2) spinal cord, with perineal neurons (Fig. 1, D and E, cyan) synapsing in the touch recipient zone (*25*) of the lateral dorsal horn (Fig. 1E and fig. S2C). Neurons innervating male prepuce (Fig. 1, D and E, yellow) projected to a

medial portion of the touch zone (Fig. 1E), whereas glans axons (Fig. 1, D and E, magenta) terminated proximal to the central canal (Fig. 1E). In females, axons from the vaginal opening targeted the medial dorsal horn, whereas those from the prepuce, which includes the clitoris, closely resembled those from the glans in males.

To visualize the peripheral anatomy of touch neurons in the perineum, we generated mice in which *Piezo2*-expressing neurons were selectively labeled by crossing a *Piezo2-Cre* allele (*26*) into a neural-specific *Snap25-LSL-GFP* reporter line (*27*). Green fluorescent protein (GFP) staining of cleared skin demonstrated that the perineum was densely innervated with lanceolate and circumferential endings surrounding hair follicles (Fig. 1F), which is consistent with innervation by a broad range of low-threshold mechanosensory neurons (LTMRs) and the PIEZO2-dependent behavioral sensitivity of mice to perineal touch.

Perineal sensory neurons exhibit high sensitivity to punctate stimulation

We developed a sacral ganglia imaging preparation to compare neural responses to a range of gentle and intense mechanical stimuli (*28*) applied to the hind-paw and perineum (fig. S3 and movie S2). Neurons innervating paw glabrous skin divided into LTMRs and high-threshold mechanosensory neurons (HTMRs) on the basis of their response selectivity (Fig. 2A; fig. S3; and supplementary materials, materials and methods). HTMRs that innervate the paw outnumbered LTMRs by a factor of 2. In particular, LTMRs exhibited graded von Frey sensitivity (Fig. 2A and fig. S3D) and could be activated by forces as low as 0.008 g, whereas HTMRs were essentially silent at forces below 0.4 g, matching behavioral withdrawal threshold and implicating HTMRs in this response. By contrast, perineal sensation was dominated by LTMRs, with ~60% of mechanosensory neurons responding to gentle stimuli (Fig. 2B and figs. S3 and S4) and broad similarity between male and female mice (fig. S4). Almost all perineal mechanosensors could be activated by von Frey stimulation (Fig. 2B, figs. S3, S4), and their calcium (GCaMP) signals were markedly stronger than for paw-innervating neurons (fig. S3D). Few HTMRs responded to the fine filaments that reliably evoked behavioral responses (Fig. 1B and movie S1). Therefore, both male and female mice are attuned to perineal LTMR input, and the stereotyped reaction to genital touch is not a sign of pain.

A broad role for PIEZO2 in perineal sensation

To measure the contribution of PIEZO2 to perigenital touch and to dissect the mechanism underlying the extreme sensitivity to von Frey stimulation, we next used the sacral imaging platform to selectively image cells that lack this stretch-gated ion channel (fig. S4D).

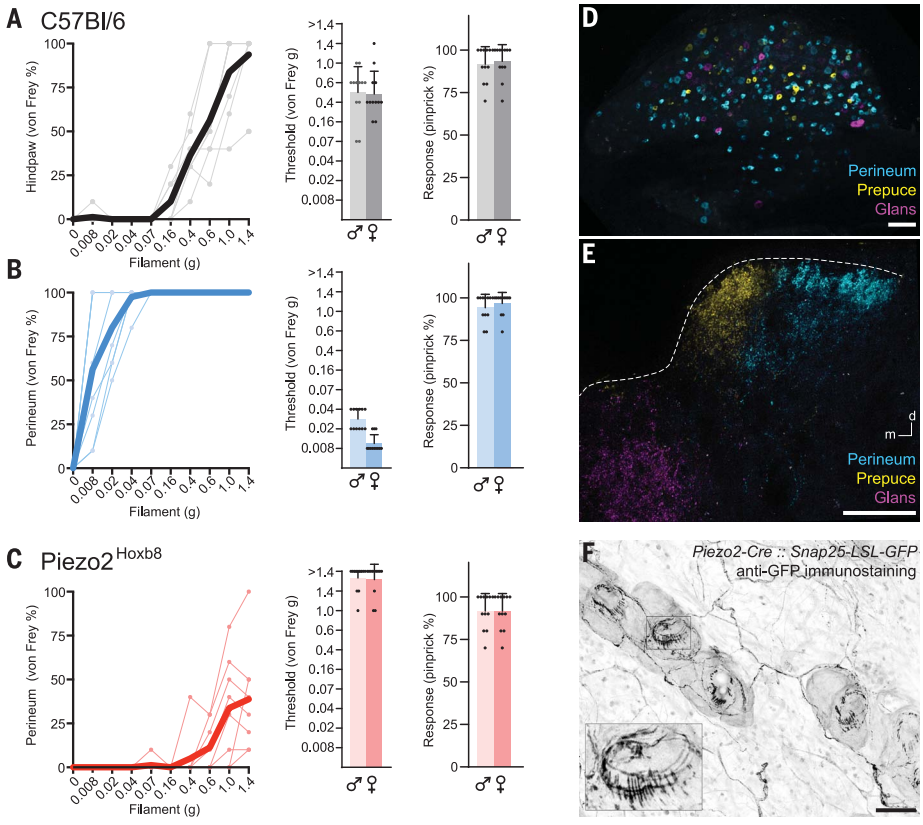


Fig. 1. Behavioral sensitivity of mice to perineal touch and underlying anatomy. (**A** to **C**) Reaction of mice to punctate touch (A) wild-type hind-paw, (B) wild-type perineum, and (C) Piezo2^{Hoxb8} perineum. (Left) Example responses for individual mice (points and thin lines; four males and four females) and mean (solid lines) to a series of calibrated von Frey filaments (grams, each tested 10 times per mouse). (Middle) Quantitation of von Frey threshold (≥5/10; n = 12 males and 12 females). Thresholds are different between all three groups [one-way analysis of variance (ANOVA) on ranks $P < 0.001$]. Wild-type females exhibited a lower perineal touch threshold than that of males (Mann-Whitney t test; $P < 0.0001$); there were no significant differences in other responses (supplementary materials, statistical reporting). (**D** and **E**) Triple-color retrograde CTB tracing from the perineum (cyan), prepuce (yellow), and glans (magenta) showing (D) cell bodies of lumbar-sacral sensory neurons in the DRG and (E) termini in the dorsal spinal cord. The dotted line indicates approximate extent of dorsal horn. In (E) and (F), n = 4 mice. Scale bars, 100 μm. (**F**) Anatomy of sensory ending of *Piezo2*-expressing sensory neurons in the perineum. (Inset) A magnified view of a single hair (boxed) highlighting prominent lanceolate and circumferential endings (n = 2 males and 1 female). Scale bar, 50 μm.

As expected, deletion of *Piezo2* (Piezo2^{cKO}) dramatically affected the mechanosensitivity of genital-innervating neurons. The great majority of responses to air puff, vibration, and brush were eliminated. Thus, mechanosensory neurons were only stimulated by pinch and were almost exclusively HTMRs (fig. S4E). The overall number of HTMRs was similar between wild-type and Piezo2^{cKO} mice (fig. S4F), which is consistent with earlier studies (*17*, *20*, *21*). Piezo2^{cKO} mice responses to von Frey stimulation were substantially reduced and recapitulated those of control perineal HTMRs (Fig. 2, C to E). These results likely explain the absence of behavioral reactions to von Frey stimulation in Piezo2^{Hoxb8} mice (Fig. 1C), support the hypothesis that perineal LTMRs drive this characteristic withdrawal in wild-type mice (movie S1), and are consistent with human reports and sensory testing (table S1).

A subset of touch neurons is required for mechanically induced erection responses

Perineal investigation and touch precedes mating in many species, including mice (*29*). These behaviors are linked to motivational drive in both partners and trigger physiological reflexes. For example, gentle retraction of the prepuce induces penile cupping (erection) and flipping (ejaculation) in spinalized rodents (*30*). We reasoned that mechanosensory input drives the erection reflex and developed an assay to monitor this in restrained awake mice. A soft, transparent tube was used to gently retract the prepuce, allowing the physiological erection reflex (extension of the penis into the tube)

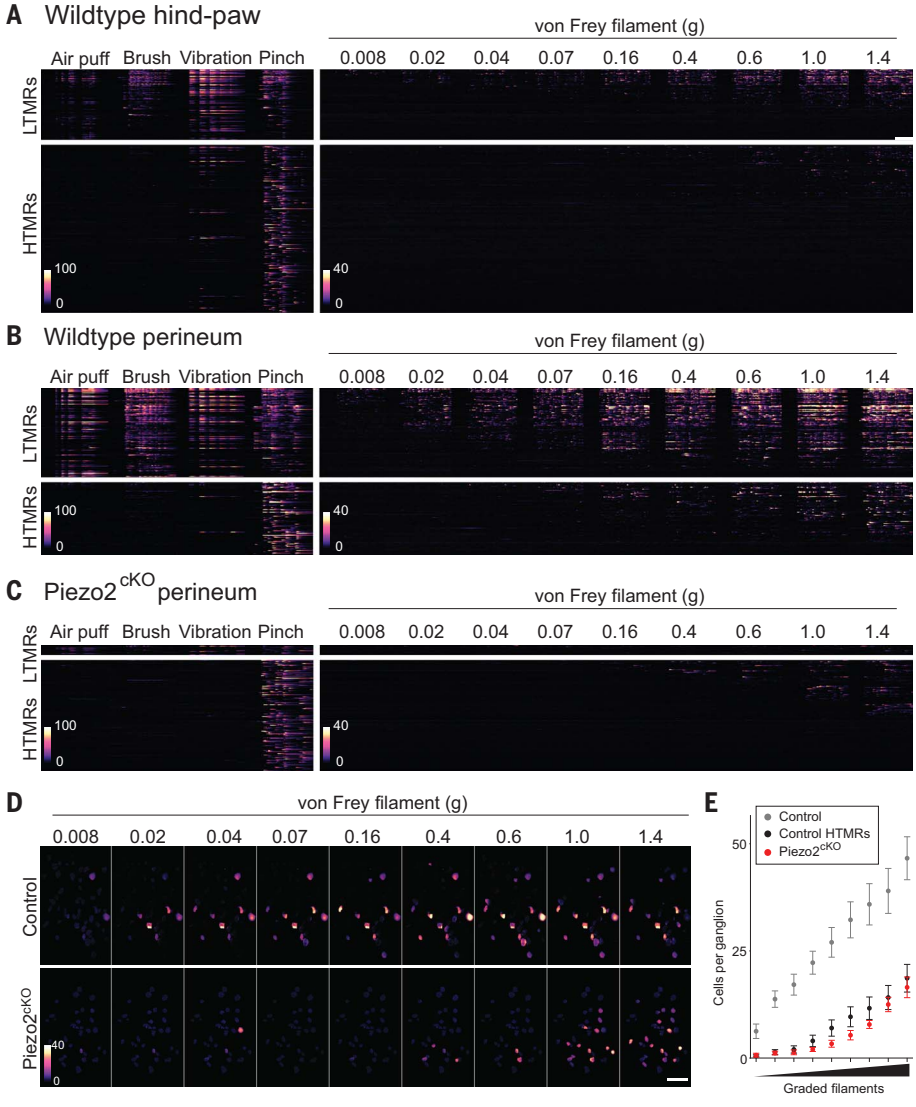


Fig. 2. Functional characterization of perineal mechanoreceptors and role of PIEZO2. (**A** to **C**) Heatmaps representing calcium (GCaMP6f) responses to (left) repetitive application of naturalistic stimuli and (right) graded von Frey stimulation. LTMRs and HTMRs are separated, and relative fluorescence changes ($\Delta F/F$) are colored as indicated. Scale bar, 10 s. (A) Wild-type hind-paw, n = 4 mice. (B) Wild-type perineum, n = 4 mice. (C) Piezo2^{cKO} perineum, n = 6 mice. Additional analysis is provided in figs. S3 and S4. (**D**) Spatial activity maps of control and Piezo2^{cKO} neurons to von Frey filaments. Scale indicates response intensity. Scale bar, 100 μm. (**E**) Quantitation of von Frey responsive neurons in control mice (gray), Piezo2^{cKO} mice (red), and response profile of control HTMRs (black) (mean ± SEM, n = 8 control mice, n = 6 Piezo2^{cKO} mice). Piezo2^{cKO} mice had fewer von Frey responsive neurons at all filament strengths (Mann Whitney U test; $P < 0.0087$).

to be scored. Wild-type controls responded in almost every single trial (Fig. 3A); isoflurane anesthesia completely eliminated responses, and local numbing of the perineum with lidocaine greatly dampened the reflex (fig. S5). As we anticipated, Piezo2^{Hoxb8} mice only very rarely exhibited penile extension in response to prepuce retraction (Fig. 3A).

Piezo2^{Hoxb8} mice exhibit broad loss of touch but also have proprioceptive (and potentially other mechanosensory) deficits (*17*, *28*). Therefore, we examined mice with more selective *Piezo2* deletions. Piezo2^{Pvalb} mice (in which *Piezo2* is inactivated by using *parvalbumin*-driven *Cre*) lack proprioceptive input but still respond to gentle touch (*20*). These mice had perfectly normal responses to prepuce retraction (Fig. 3A) despite severe ataxia. We also generated *Piezo2* deletions using an *Scn10a-Cre* line, which is commonly used to target a broad range of nociceptors, including HTMRs (*31*). Perineal HTMR responses are PIEZO2 independent (Fig. 2 and fig. S4); therefore, these mice (Piezo2^{Scn10a}) were predicted to have normal proprioception, touch, and consequently erection reflexes. Piezo2^{Scn10a} mice walked with normal gait, and recombination of *Scn10a-Cre* in sacral ganglia neurons was faithful (Fig. 3, B and C, and fig. S6A), with only a few large-diameter *Scn10a*-negative LTMRs labeled (fig. S6A). Nonetheless, Piezo2^{Scn10a} mice displayed severe deficits in their erection reflex, closely recapitulating the phenotype of Piezo2^{Hoxb8} animals (Fig. 3A) and the effects of lidocaine (fig. S5A). Single-cell sequencing data from lumbar DRG (*32*) and trigeminal neurons (*33*) validate *Scn10a* as a robust marker for nociceptors but reveal expression in c-fiber LTMRs (cLTMRs). We used in situ hybridization (ISH) to confirm coexpression of *Scn10a*, the cLTMR marker *Tyrosine hydroxylase* (*Th*) (*24*), and *Piezo2* in sacral ganglia (Fig. 3C), with only very limited recombination in other potential LTMRs (fig. S6A). Because cLTMR responses to gentle mechanical stimulation depend on *Piezo2* expression (*34*), these data strongly suggest a causal role for perineal cLTMR input in triggering the erection reflex. Consistent with this hypothesis, tdTomato–positive lanceolate endings (typical of cLTMRs) surround perineal hair follicles in *Scn10a-Cre*, *Ai9* mice (Fig. 3D). Moreover, functional imaging of perineal touch responses in *Scn10a-Cre*, *Ai95* mice revealed that neurons responding to gentle mechanical stimuli (fig. S6, B and C) had uniform small diameters, as would be expected for cLTMRs (*24*, *34*).

Severely impaired sexual function in mice lacking PIEZO2

Loss of a touch-induced erection response in Piezo2^{Hoxb8} males should impair mating. Indeed, 10 pairs of mating-age Piezo2^{Hoxb8} males and females housed together for 6 months never produced pups, whereas wild-type (C57Bl/6) controls delivered 61 litters in this time (range, five to seven litters per pair). To assess copulatory success more directly, we also examined the frequency of vaginal plug formation after introducing virgin females in estrus to single housed males; to eliminate bias from prior experience, all mice were naïve. For C57Bl/6 mice, 7 from 10 homozygous pairings had plugs after 4 hours (Fig. 3E). By contrast, plugs were never seen for Piezo2^{Hoxb8} male mice when paired either with Piezo2^{Hoxb8} or wild-type females (Fig. 3E). As predicted from their normal erection reflexes, Piezo2^{Pvalb} males successfully mated with C57Bl/6 females despite severe ataxia (Fig. 3E). However, Piezo2^{Scn10a} males failed to plug receptive C57Bl/6 females, substantiating the importance of PIEZO2-dependent mechanosensory input for male mating behavior (Fig. 3E). Although loss of erection reflexes may explain why Piezo2^{Hoxb8} mice fail to breed, mechanosensation probably has additional

roles in mating. For example, female mice have similar PIEZO2-dependent perineal mechanosensitivity to males (fig. S4) and are even more sensitive to perigenital touch (Fig. 1); Piezo2^{Hoxb8} females exhibited strong mating deficits when paired with wild-type males: 9 from 10 remained unplugged after 4 hours (Fig. 3E).

Ethogram analysis of female intruder assays (Fig. 3F) assess motivation by quantifying stereotyped male behaviors, including partner-grooming, anogenital chemosensory investigation, and mounting attempts (*35*, *36*). We analyzed behavior for 1 hour after introduction of receptive females (supplementary materials, materials and methods). Control animals exhibited considerable variation in mating behavior (Fig. 3F) but in every case (n = 10 pairs of mice) engaged in chemosensory investigation and mounting attempts shortly after introduction of the female. Similarly, pairs of Piezo2^{Hoxb8} males and females (n = 10 pairs) (Fig. 3F) as well as male or female Piezo2^{Hoxb8} mice paired with C57Bl/6 partners (n = 10 pairs in each case) (fig. S5B) exhibited strong sexually motivated behavior, not very different from controls. However, Piezo2^{Hoxb8} males never achieved intromission, which was regularly observed in wild-type controls. Similarly, Piezo2^{Scn10a} males paired with receptive C57Bl/6 females showed normal sexual motivation (n = 10 males and 10 females) (fig. S5B) but without copulatory success (Fig. 3E). Moreover, Piezo2^{Hoxb8} females paired with C57Bl/6 males also engaged in premating behavior, including male mounting attempts (fig. S5B), but Piezo2^{Hoxb8} females adopted a sit-rejection posture, preventing intromission (*37*). These data show that mechanosensation plays a substantial role in productive mating and exposes dimorphic need for PIEZO2 and gentle touch in sexual function.

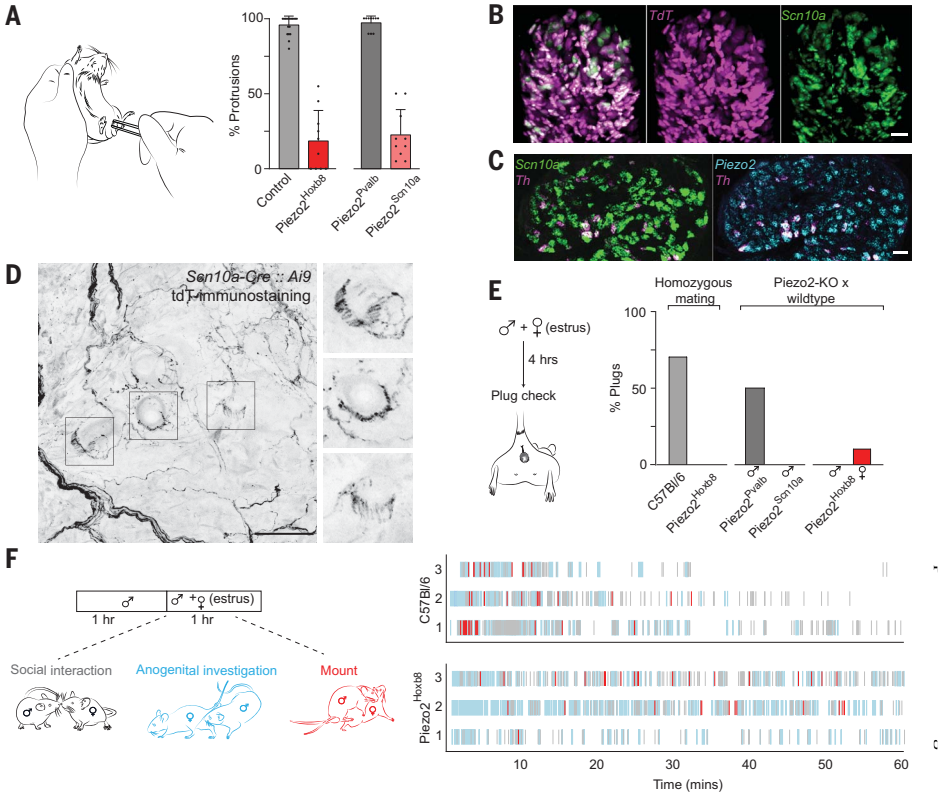


Fig. 3. A role for PIEZO2-dependent perineal mechanosensation in mating. (**A**) Physiological responses of male mice to perineal stimulation with transparent soft tubing. Penile protrusion was scored for two sets of 10 trials. Bars indicate mean ± SEM, and points indicate individual responses. Control versus Piezo2^{Hoxb8} mice and Piezo2^{Pvalb} versus Piezo2^{Scn10a} mice were different (Mann-Whitney U test; $P < 0.0001$; n = 18 control mice; n = 10 *Piezo2*-deleted mice). (**B**) Representative whole-mount ISH of sacral ganglion showing faithful recombination [*tdTomato* (*TdT*); magenta] of *Scn10a-Cre* mouse in *Scn10a* (green) neurons; >90% (670 of 738) *TdT* cells expressed *Scn10a* (n = 3 ganglia). Scale bar, 50 μm. (**C**) Example ISH of sacral ganglion section probed for *Th* (magenta), *Scn10a* (green), and *Piezo2* (cyan), illustrating expression of *Scn10a* and *Piezo2* in cLTMRs identified with *Th* (n = 6 ganglia). Scale bar, 50 μm. (**D**) Anatomy of sensory ending of *Scn10a*-expressing sensory neurons in the perineum (maximum projection, full-stack). (Right) Magnified and focal views of single hairs (boxed at left), highlighting lanceolate endings of *Scn10a-Cre*–labeled neurons (n = 2 male mice). (**E**) Successful mating scored by vaginal plugs (n = 10 mice). Differences are significant for C57Bl/6 versus Piezo2^{Hoxb8} mice (P = 0.0031) and Piezo2^{Pvalb} versus Piezo2^{Scn10a} mice (P = 0.0325) (Fisher's exact test, two-tailed). (Right) Mating success for female and male Piezo2^{Hoxb8} mice with C57Bl/6 partners. (**F**) Representative ethogram plots showing sexual motivation of three isogenic pairings of C57Bl/6 and 3 Piezo2^{Hoxb8} mice: social interaction (gray), anogenital investigation (pale blue), and mounting attempts (red).

Impact of PIEZO2 in human sexual experience

The genital sensation of a man with complete loss of PIEZO2 function and comprehensive touch- and proprioception-related studies of individuals with PIEZO2-deficiency syndrome (*11*, *21*) demonstrate strongly conserved roles for PIEZO2 in mammalian mechanosensation. For humans, sexual experience is not simply related to reproduction but is central to large parts of many people's social lives and behavior. Information from human clinical evaluations (n = 5; three men and two women) (table S1) provided several consistent themes about the role of gentle touch in sex. First, these individuals with biallelic loss of function (table S1A) had diagnostic clinical presentation, with loss of proprioception, absent vibration sensing, highly elevated touch threshold, and scoliosis but no cognitive difficulties, and all underwent puberty without clinically relevant problems. Second, all five people with PIEZO2 deficiency reported being sexually active and able to be aroused by physical genital stimulation, erotic thoughts, or videos, reflecting motivation seen in Piezo2^{Hoxb8} mice (Fig. 3F). Third, individuals with PIEZO2 deficiency reported delayed, attenuated, or absent physiological responses to gentle genital stimulation. This included clinical diagnosis of hypo-orgasmia for the male and anorgasmia for the female participants, which again is consistent with the animal model. However, the five people had strategies to compensate for deficits in genital sensation (table S1B).

Discussion

Erogenous touch conveys different meanings according to circumstance; however, many key details remain unknown. We explored how deficits in PIEZO2-dependent mechanosensation interfere with perigenital sensation, physiological response, copulation, and reproduction. Our results demonstrate that PIEZO2-dependent touch is required for all of these in mice.

Anatomical studies have identified specialized corpuscles composed of myelinated afferents likely involved in genital sensation (*38*, *39*). Our data strongly implicate an additional type of touch neuron, the perineal cLTMRs, as crucial drivers of sexual function. Previous studies in mice and humans suggest specialized roles for cLTMRs in conveying affective and pleasurable touch (*40*, *41*). Thus, it is of note that five

individuals without PIEZO2 function described sexual activity as satisfying and rewarding despite marked mechanosensory deficits and clinical evaluations of hypo-orgasmia (men) and anorgasmia (women). We have previously shown that for humans, other types of sensory input can compensate for deficits caused by loss of PIEZO2 function (*11*). For example, these individuals use vision to overcome proprioceptive deficits and mechanonociception or thermosensation to mitigate deficits in touch (*11*). This is also true for human sexual touch (table S1). Nonetheless, the crucial role of PIEZO2 for perineal touch in mice and humans may have therapeutic potential: Topical PIEZO2 inhibitors could provide targeted relief of genital hypersensitivity and pain, whereas agonists of PIEZO2 are candidates for alleviating genital hyposensitivity.

There are a number of limitations to this work. For example, PIEZO2 deficiency is extremely rare, and we were unable to carry out detailed quantitative sensory testing in a larger group of human subjects. Additionally, functional imaging experiments were carried out in anesthetized mice, precluding evaluation of responses during mating. Moreover, although we showed the necessity of gentle touch input for mating, we have not yet demonstrated the sufficiency of this sensory pathway for sexual function in awake behaving animals. We also anticipate that there are likely to be additional specialized roles for mechanosensory neurons in mating that were not revealed in this study.

Even the very gentlest of perineal touches elicits a highly stereotyped startle reaction from mice that is easy to anthropomorphize (movie S1). This PIEZO2-dependent response is quite different from touch to other parts of the body, which typically evokes more modest reactions and does so only at much greater forces. PIEZO2-dependent perineal touch is also a crucial driver of successful mating both for male and female mice. Future studies should help define additional subtypes of sensory neurons needed for sexually dimorphic reactions and how perigenital sensation is organized in the spinal cord and brain to prioritize salience. Ultimately, however, the profound impact of PIEZO2 deficiency that we describe provides a sensory basis at the molecular and cellular level for an aspect of life that throughout history has engaged in human imagination (*42*) and thought (*1*).

REFERENCES AND NOTES

1. C. Darwin, *The Descent of Man and Selection in Relation to Sex* (Murray, 1871), vols. 1 and 2.
2. J. J. Dinsmore, *Auk* **87**, 305–321 (1970).
3. S. J. Gould, *Evolution* **28**, 191–220 (1974).
4. M. A. Changizi, Q. Zhang, S. Shimojo, *Biol. Lett.* **2**, 217–221 (2006).
5. B. D. Charlton, M. S. Martin-Wintle, M. A. Owen, H. Zhang, R. R. Swaisgood, *R. Soc. Open Sci.* **5**, 181323 (2018).
6. V. S. Mandiyan, J. K. Coats, N. M. Shah, *Nat. Neurosci.* **8**, 1660–1662 (2005).
7. L. Stowers, T. E. Holy, M. Meister, C. Dulac, G. Koentges, *Science* **295**, 1493–1500 (2002).
8. C. C. Voigt *et al*., *J. Mammal.* **89**, 1401–1410 (2008).
9. C. A. Kell, K. von Kriegstein, A. Rösler, A. Kleinschmidt, H. Laufs, *J. Neurosci.* **25**, 5984–5987 (2005).
10. B. Coste *et al*., *Science* **330**, 55–60 (2010).
11. A. T. Chesler *et al*., *N. Engl. J. Med.* **375**, 1355–1364 (2016).
12. S. S. Ranade *et al*., *Nature* **516**, 121–125 (2014).
13. Y. García-Mesa *et al*., *J. Anat.* **238**, 446–454 (2021).
14. Y. García-Mesa *et al*., *J. Anat.* **239**, 892–902 (2021).
15. I. Abdus-Saboor *et al*., *Cell Rep.* **28**, 1623–1634.e4 (2019).
16. T. Hua *et al*., *Nat. Neurosci.* **23**, 854–868 (2020).
17. S. E. Murthy *et al*., *Sci. Transl. Med.* **10**, eaat9897 (2018).
18. V. E. Abraira, D. D. Ginty, *Neuron* **79**, 618–639 (2013).
19. K. Nonomura *et al*., *Nature* **541**, 176–181 (2017).
20. S.-H. Woo *et al*., *Nat. Neurosci.* **18**, 1756–1762 (2015).
21. M. Szczot *et al*., *Sci. Transl. Med.* **10**, eaat9892 (2018).
22. M. L. Sorrells *et al*., *BJU Int.* **99**, 864–869 (2007).
23. C. B. Bleustein, H. Eckholdt, J. C. Arezzo, A. Melman, *J. Urol.* **169**, 2266–2269 (2003).
24. L. Li *et al*., *Cell* **147**, 1615–1627 (2011).
25. V. E. Abraira *et al*., *Cell* **168**, 295–310.e19 (2017).
26. S. H. Woo *et al*., *Nature* **509**, 622–626 (2014).
27. L. Madisen *et al*., *Neuron* **85**, 942–958 (2015).
28. K. L. Marshall *et al*., *Nature* **588**, 290–295 (2020).
29. F. A. Beach, *Horm. Behav.* **7**, 105–138 (1976).
30. B. D. Sachs, *Physiol. Behav.* **24**, 489–492 (1980).
31. N. Agarwal, S. Offermanns, R. Kuner, *Genesis* **38**, 122–129 (2004).
32. N. Sharma *et al*., *Nature* **577**, 392–398 (2020).
33. M. Q. Nguyen, C. E. Le Pichon, N. Ryba, *eLife* **8**, e49679 (2019).
34. L. J. von Buchholtz *et al*., *Neuron* **109**, 285–298.e5 (2021).
35. C. J. Burnett *et al*., *eLife* **8**, e44527 (2019).
36. S. X. Zhang *et al*., *Nature* **597**, 245–249 (2021).
37. X. Zhou *et al*., *Science* **379**, 820–825 (2023).
38. R. D. Johnson, Z. Halata, *J. Comp. Neurol.* **312**, 299–310 (1991).
39. D. Ohmori, *Anat. Embryol.* **70**, 347–410 (1924).
40. D. Huzard *et al*., *Sci. Adv.* **8**, eabo7566 (2022).
41. H. Olausson *et al*., *Nat. Neurosci.* **5**, 900–904 (2002).
42. M. Vatsyayana, *The Kama Sutra of Vatsyayana*, R. F. Burton, B. Indrajit, S. P. Bhide, Transl. (1883).

ACKNOWLEDGMENTS

We thank T. Delong, M. Singh, and M. Bradson for distributing questionnaires to human subjects and collating clinical notes; members of the clinical staff at NCCIH and Bönnemann groups for support in arranging human studies; and M. Szczot, N. Ghitani, D. Macdonald, and R. Hardy for input and expertise with animal experiments. Hormone analysis was carried out by the University of Virginia Center for Research in Reproduction Ligand Assay and Analysis Core. **Funding:** This work was supported by the National Institutes of Health, NCCIH Z01-ZIAAT000028 (to A.T.C.); National Institutes of Health, NIDCR Z01-ZIADE000561 (to N.J.P.R.); National Institutes of Health, NINDS Z01-ZIANS003129 (to C.G.B.); and the Howard Hughes Medical Institute (to A.P.)
Author contributions: Conceptualization: R.M.L., N.J.P.R., and A.T.C. Methodology: R.M.L., M.F., L.J.v.B., E.F., C.G.B., N.J.P.R., and A.T.C. Investigation: R.M.L., M.F., J.O., E.F., M.R.S.-V., M.N., M.Q.N., M.J., and D.S. Funding acquisition: A.P., C.G.B., N.J.P.R., and A.T.C. Supervision: A.P., C.G.B., N.J.P.R., and A.T.C. Writing – original draft: R.M.L., L.J.v.B., N.J.P.R., and A.T.C. Writing – review and editing: R.M.L., M.F., L.J.v.B., E.F., A.P., C.G.B., N.J.P.R., and A.T.C.
Competing interests: The authors declare that they have no competing interests. **Data and materials availability:** All data generated and/or analyzed during the current study are provided in the supplementary materials and/or have been deposited in Dryad. **License information:** Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse. This article is subject to HHMI's Open Access to Publications policy. HHMI lab heads have previously granted a nonexclusive CC BY 4.0 license to the public and a sublicensable license to HHMI in their research articles. Pursuant to those licenses, the author-accepted manuscript (AAM) of this article can be made freely available under a CC BY 4.0 license immediately upon publication.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adg0144
Materials and Methods
Figs. S1 to S7
Table S1
References (*43*–*46*)
Movies S1 and S2

Submitted 27 November 2022; accepted 13 July 2023
10.1126/science.adg0144

By **Zheng Li**

Small does not mean less

During my graduate studies, countless seminars taught me one thing: To make it in science, one must design and carry out big projects, with the aim of addressing fundamental questions and making major breakthroughs. As I prepared to establish my own research lab, advice from a senior colleague reinforced this notion: "Try everything you can to do a big project! Once you have a high-impact publication, everything in the future will be much smoother, and then your career is settled." So, when I landed an associate professor position, my goal seemed obvious: I should devise and execute big projects. I thought this was the only way to lay a foundation for a successful academic research career. The path, however, would not be as clear as I had envisioned.

When I contacted prominent researchers to suggest collaborating, as is typically required for high-profile projects, I was often turned down. They frequently replied enthusiastically to start, agreeing my proposed project sounded interesting and promising. But when they learned I had just established my lab, they became skeptical and hesitant, ultimately saying no. And when I sought grant funding, my lack of experience and reputation was cited as a reason to deny support for my ambitious proposals. I also had to admit that the prospect of administering a big project felt a bit daunting, given the responsibilities of establishing the infrastructure for my new lab and training my eager yet inexperienced students.

"I began to ask myself: Were big projects really the only route?"

Those students also struggled with my single-minded ambition to execute a big project. They told me that, given how much I had invested in it financially, intellectually, and professionally, the pressure was too much and they worried they would not be able to accomplish what was needed.

I spent about 2 years stuck in this mindset, excited about my big ideas but making little actual progress, leading to frustration and depression. I knew I couldn't go on this way. As I neared my breaking point, I began to ask myself: Were big projects really the only route? Are small projects really such a terrible thing?

When I raised the topic with senior colleagues, they continued to assure me that big was the way to go. Nonetheless, my adventurous spirit finally won out, and I began to conceive a number of exploratory investigations. The scientific questions they pursued were not necessarily significant, but I enjoyed following my curiosity.

This pivot freed my creativity. I went back to square one, designing projects from scratch rather than just mimicking ideas from high-impact papers. When I shook off the feeling that I must publish exclusively in high-profile journals, I became brave enough to step into fields I was not familiar with. By reading papers in new areas and even borrowing textbooks from the library, I found novel connections and interdisciplinary ideas to pursue. When I discussed these adventurous, low-cost (both financially and intellectually) projects with colleagues, mostly early-career researchers, they were keen to collaborate. We secured funding, and my students were energized. When I published papers out of these projects, they were not exciting enough to make headlines—but for me, it was a moment of joy and self-contentment.

It has now been 3 years since I embraced this "small" approach. Sometimes senior colleagues question my research program, worrying I lack focus and ambition as I explore many ideas heading in different directions. Some say I'm wasting my time with projects that aren't important enough. But I disagree. When I do these studies, I also find clues for future big projects—just in time for the infrastructure in and around my lab to be ready for such work. And I'm excited to pursue these more profound questions. But I also know that they are not the be-all, end-all of scientific research, and that there is much to be said for small projects and the freedom and creativity they can foster. Going forward, I will still commit a considerable part of my time to exploring small projects. Small, I've learned, does not mean less. ■

Zheng Li is an associate professor at Yunnan University. Send your career story to SciCareerEditor@aaas.org.